ON THE NUMBER OF IRREDUCIBLE FACTORS OF A POLYNOMIAL

by

CHRISTOPHER G. PINNER, B.Sc., M.Sc.

DISSERTATION
Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN
August, 1992
ON THE NUMBER OF IRREDUCIBLE FACTORS OF A POLYNOMIAL

APPROVED BY
SUPERVISORY COMMITTEE:

Supervisor:  
Jeff Caela
Rudy Heath Brown
Frank Smith III
David Nash
Samuel White
To my parents
Acknowledgments

I would particularly like to thank my supervisor Jeff Vaaler, whose invaluable insights, ready advice, and exceptional generosity with his time, were always greatly appreciated.

Thanks should go as well to the members of the dissertation committee for their careful detection of flaws in the manuscript. I must also mention here Roger Heath-Brown and Maurice Dodson, who first encouraged my interest in number theory.

Finally, I am extremely grateful to the Department of Mathematics and to The University of Texas for their sustained and generous financial support.

CHRISTOPHER G. PINNER

The University of Texas at Austin
August, 1992.
ABSTRACT

ON THE NUMBER OF IRREDUCIBLE FACTORS OF A POLYNOMIAL

by

CHRISTOPHER G. PINNER, B.Sc., M.Sc.

SUPERVISING PROFESSOR: JEFFREY VAALER

Given a polynomial $F(x)$ with coefficients in an algebraic number field $k$ we consider the problem of estimating the number of irreducible factors of $F(x)$ over $k[x]$; counting these factors both with and without multiplicity. The cyclotomic and non-cyclotomic factors are treated separately.

A bound on the total number of the cyclotomic factors conjectured earlier by Schinzel is verified. New estimates are given for the number of $n$-th and primitive $n$-th roots of unity amongst the roots of $F(x)$. These bounds are all expressed in terms of the degree and ‘height’ of the polynomial together with a fairly explicit dependance upon the particular field $k$. A second set of results
depending upon the degree, field, and number of non-zero coefficients of $F(x)$ (rather than the height) are also presented in the cyclotomic case. Examples are constructed showing the essential sharpness of some of the inequalities.

Finally, an equivalence is demonstrated between certain upper bounds on the number of non-cyclotomic factors of a polynomial in $k[x]$ and Lehmer’s conjectured lower bound on the Mahler measure of a non-cyclotomic polynomial in $k[x]$. 
# Table of Contents

Acknowledgments iv

ABSTRACT v

Table of Contents vii

1. A review of the results 1
   1.1 Introduction 1
   1.2 The statement of the various theorems 7
   1.3 A few brief comments on the theorems 11

2. Basic definitions and inequalities 16
   2.1 Preliminaries on valuations 16
   2.2 Heights and measures on polynomials 23
   2.3 The operator $\mathcal{L}$ and some of its properties 35

3. The non-cyclotomic factors 46
   3.1 Proof of theorem 1 46
   3.2 Non-cyclotomic bounds and Lehmer’s conjecture 49

4. The cyclotomic factors and theorem 2 56
   4.1 The key inequalities 56
   4.2 The constant $c(k)$ and related sums 61
   4.3 A generalisation of theorem 2 67
Chapter 1

A review of the results

1.1 Introduction

Let \( k \) be an algebraic number field and let \( F(x) \) be a polynomial in \( k[x] \) of degree \( \partial(F) \) with \( F(0) \neq 0 \). We shall consider here the problem of bounding the number of irreducible factors of \( F(x) \) in \( k[x] \), counting these factors both with and without multiplicity. In the special case \( k = \mathbb{Q} \) both questions have been previously studied by Schinzel [21] and [22] and by Dobrowolski [9]. Of course we always have the trivial bound \( \partial(F) \) for such quantities and in generality this is plainly all we can say. However if the ‘height’ \( H(F) \) or the number of non-zero coefficients \( N(F) \) of \( F(x) \) is ‘small’ we should expect to do a bit better. Here \( H(F) \) is the absolute height on the vector of coefficients of \( F(x) \), to be defined precisely in section 2.2. When \( F(x) \) is in \( \mathbb{Z}[x] \) we note that \( H(F)^2 \) is just the sum of the squares of the coefficients of \( F(x) \) (after removing any common factors from them).

In fact it will often be convenient to use the different height \( \nu(F) \) defined in section 2.2. For a polynomial \( F(x) \) in \( \mathbb{Z}[x] \), \( \nu(F) \) is simply the supremum of \( |F(x)| \) on the unit circle, after again removing any common factors from the coefficients. In view of the inequality

\[
\log H(F) \leq \log \nu(F) \leq 2 \log H(F)
\]
proved in (2.22) these two heights will be virtually interchangeable in most of our results.

It is plain from the example $x^n - 1$ that we cannot hope for a bound depending solely upon $H(F)$ or $N(F)$. However, the factors that occur here are all cyclotomic. So it is not surprising that we are led naturally by the methods employed to consider the cyclotomic and non-cyclotomic factors separately.

We shall suppose that the familiar monic, irreducible $n$-th cyclotomic polynomials $\Phi_n(x)$ of $\mathbb{Z}[x]$ split up into the monic irreducible cyclotomic polynomials $\Phi_{n,s}(x)$ of $k[x]$ as follows:

$$\Phi_n(x) = \prod_{s=1}^{\delta(k;n)} \Phi_{n,s}(x). \quad (1.1)$$

If $\zeta_n$ denotes a primitive $n$-th root of unity then, as $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ is a galois extension, we have

$$\partial(\Phi_{n,s}) = [k(\zeta_n) : k] = \frac{\phi(n)}{[k \cap \mathbb{Q}(\zeta_n) : \mathbb{Q}]}.$$

(1.2)

It follows at once that

$$\delta(k; n) = [k \cap \mathbb{Q}(\zeta_n) : \mathbb{Q}] \leq [k' : \mathbb{Q}] \quad (1.3)$$

where $k'$ denotes the maximal abelian subfield of $k$. We shall suppose throughout that the factorisation of $F(x)$ takes the form

$$F(x) = x^L \left\{ \prod_{n=1}^{\infty} \prod_{s=1}^{\delta(k;n)} \Phi_{n,s}(x)^{\varepsilon(n,s)} \right\} \left\{ \prod_{i=1}^{I} f_i(x)^{m(i)} \right\} \quad (1.4)$$

where the $f_i(x) \neq x$ are distinct irreducible non-cyclotomic polynomials in $k[x]$. We shall use $\xi_i$ to denote a generic root of $f_i(x)$. 

As with Schinzel and Dobrowolski our bounds on the the number of non-cyclotomic factors follow from lower bounds on the Mahler measure of a non-cyclotomic polynomial in $\mathbb{Q}[x]$. The Mahler measure $\mu(F)$ will be defined precisely in section 2.2, although for a polynomial in $\mathbb{Z}[x]$ it is just the absolute value of the product of its roots outside the unit circle and its leading coefficient (after removing any common factors from the coefficients). It is worth noting that the truth of Lehmer's conjectured lower bound for the Mahler measure of a non-cyclotomic polynomial would in fact lead us to a bound on the number of non-cyclotomic factors of $F(x)$ that did depend solely upon $H(F)$. Conversely in section 3.2 we show how such an upper bound on the number of non-cyclotomic factors (or the maximal multiplicity of such a factor) would actually be equivalent to Lehmer's Conjecture. Use, as here, of Dobrowolski's lower bound produces instead an upper bound which does depend slightly upon the degree as well as $\log H(F)$. The form still however well reflects the fact that a polynomial of 'small' height really cannot have too many non-cyclotomic factors. Although not to the same extent as with the non-cyclotomic factors, it certainly seems plausible that the smaller the height the fewer total factors a polynomial can have. In view of this we set

$$r = r(F) = \max \left\{ 3, \frac{\partial(F)}{\log \nu(F)} \right\} \quad (1.5)$$

where we would expect our bounds to improve with increasing $r$.

It was observed by Schinzel that knowledge of the resultant of two cyclotomic polynomials in $\mathbb{Z}[x]$ could be used to produce a bound on the total number of cyclotomic factors of a polynomial in $\mathbb{Z}[x]$. In order to deal with an arbitrary number field we introduce the semi-resultant resembling operator $\mathcal{L}$.
on pairs of polynomials in $\mathbb{Q}[x]$. We define $\mathcal{L}$ precisely in section 2.3 although in the case of two monic polynomials $G_1(x)$ and $G_2(x)$ in $\mathbb{Z}[x]$ with no common zeros we note that $\mathcal{L}(G_1, G_2) = \log |\text{resultant}(G_1, G_2)|$. In general, information about both the original polynomials and their conjugates over $\mathbb{Q}[x]$ is encapsulated in $\mathcal{L}$ by summing over the places of a suitable field. In particular the property $\mathcal{L}(G_1, G_2) \geq 0$ for all $G_1(x), G_2(x)$ in $\mathbb{Q}[x]$ will replace the observation (trivial, but essential to Schinzel's original argument) that for two polynomials $G_1(x), G_2(x)$ in $\mathbb{Z}[x]$ (with non-zero resultant) $\log |\text{resultant}(G_1, G_2)| \geq 0$. By considering $\mathcal{L}(\Phi_n, F)$ for each cyclotomic polynomial $\Phi_n(x)$ we gain a set of inequalities bounding the multiplicities $e(m, s)$ (where either $m/n$ or $n/m$ is a prime power) in terms of the multiplicities $e(n, s)$ and height $\nu(F)$. Arithmetical manipulation of these inequalities leads to our improved bound on the total number of cyclotomic factors of $F(x)$, and also to the seemingly new bounds on the number of $n$-th and primitive $n$-th roots of unity amongst the roots of $F(x)$. We also give a bound on the sum of the 'average' multiplicity of an $n$-th cyclotomic factor

$$a(n) = \frac{1}{\delta(k; n)} \sum_{s=1}^{\delta(k,n)} e(n, s)$$

(1.6) – such weighting removes any field dependence from the problem. The bounds produced in theorem 2 and 3 depend solely upon $\partial(F)$, $r$ and the field $k$, becoming non-trivial once $r \gg_{e, k} 1$. As expected this improvement over the trivial bound does become more pronounced as $r$ increases. Examples in section 4.6 will demonstrate how sharp these estimates can be. We should note that all these bounds are never any smaller than $\sqrt{\partial(F)} \log \partial(F)$. However if these same factors are simply counted without their multiplicities $e(n, s)$, bounds of
size $\sqrt{\vartheta(F)}$ are fairly easily obtained. In theorem 4 we give several results of this general type.

In theorem 5 we consider further bounds on the number of cyclotomic factors which this time involve the number of non-zero coefficients $N(F)$ of $F(x)$ rather than the height $\nu(F)$ or $H(F)$. The example $x^n - 1$ again shows that some other parameter (reflecting for example divisors of the exponents) has to be incorporated into such a bound. To this end if

$$F(x) = \sum_{i=1}^{N} a_i x^{n_i}$$

with the $a_i \in k$ ($a_1 \neq 0$) and $0 = n_1 < n_2 < \ldots < n_N$ we define the complicated expression

$$\Delta(F) = \# \{d \in G : d \mid n_i - n_j \text{ some } 1 \leq j < i \leq N\} \prod_{\substack{p \mid E \\; \text{and} \; p \leq N}} \left(1 + \frac{1}{p}\right) \quad (1.7)$$

where

$$E = n_2 n_3 \ldots n_N$$

and

$$G = \left\{d \mid E : d = RS \text{ with } R \mid n_i \text{ some } i \geq 2 \text{ and } S \mid \prod_{p \leq N} p \right\}. \quad (1.8)$$

Since often such detailed knowledge of the exponents is unavailable we note the rough bounds

$$\Delta(F) \ll \left(\sum_{i=1}^{N(F)} \sum_{j-i+1}^{N(F)} d(n_j - n_i)\right) \log N(F) \ll \vartheta(F)^c N(F)^2 \quad (1.9)$$

where $d(n)$ is the classical divisor function.
It was shown by Mann [18] that a primitive \( m \)-th root of unity \( \zeta_m \) cannot be a root of the polynomial \( F(x) \) in \( \mathbb{Z}[x] \) unless \( m \) is of the form \( m = m_1m_2 \) where \( m_1 \) divides at least one of the exponents \( n_i \) of \( F(x) \) and \( m_2 \) is squarefree consisting solely of primes less than \( N(F) \). In chapter 5 we develop a version of Mann's theorem which not only allows us to consider coefficients in an arbitrary number field, but which contains much more information on the shape of the \( m \)'s that can occur. Crudely counting all such possibilities gives us the bound in theorem 5.

In most of our results the field dependence is fairly simple. In several bounds the field constants \( c(k) \) and \( b(k) \) occur where

\[
c(k) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{\infty} \delta(k; n) \quad [k(\zeta_n):k] \leq N
\]

\[
b(k) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{\infty} 1. \quad [k(\zeta_n):k] \leq N
\]

Evaluating such expressions leads us naturally to consider the field constant

\[
J = J(k) = \min \{ j \geq 1 : k' \subseteq \mathbb{Q}(\zeta_j) \}
\]

- the existence of such a \( J \) being guaranteed by the Kronecker-Weber theorem. Since \( \mathbb{Q}(\zeta_n) \cap \mathbb{Q}(\zeta_J) = \mathbb{Q}(\zeta_{\gcd(n,J)}) \) we see that in fact

\[
\delta(k; n) = \delta(k; \gcd(n, J)).
\]

In section 4.2 we show the following equivalent formulations of \( c(k) \) and \( b(k) \)

\[
c(k) = \prod_{p \mid J} \left( 1 + \frac{1}{p(p-1)} \right) \sum_{m \mid J} \delta(k; m)^2 \phi(J/m) \frac{\phi(m)(J/m)}{\phi(m)}
\]

\[
b(k) = \prod_{p \mid J} \left( 1 + \frac{1}{p(p-1)} \right) \sum_{m \mid J} \delta(k; m) \phi(J/m) \frac{\phi(m)(J/m)}{\phi(m)}
\]
together with the more tangible upper bounds

\[
c(k) \leq \frac{\zeta(2)\zeta(3)}{\zeta(6)} \min \left\{ [k' : Q]^2, \ d(J)[k' : Q], \ J \right\} \\
b(k) \leq \frac{\zeta(2)\zeta(3)}{\zeta(6)} \min \left\{ d(J), \ [k' : Q] \right\} .
\]

Finally we note that the constants implied by the Vinogradov \( \ll \) and \( O \) symbols are everywhere absolute and computable (in particular they do not depend upon the field \( k \) unless specified by a suffix e.g. \( \ll_k \)).

1.2 The statement of the various theorems

We obtain the following inequalities for the number of non-cyclotomic factors of a polynomial:

**Theorem 1** Let \( F(x) \) be a polynomial in \( k[x] \) factoring into irreducibles in \( k[x] \) as shown in (1.4). Then for any \( \varepsilon > 0 \)

\[
\sum_{i=1}^{I} m(i)[k(\xi_i) : Q(\xi_i)] \leq (1 + \varepsilon)^4 |k : Q| \partial(F) \frac{1}{r} \left( \frac{\log r}{\log \log r} \right)^3
\]

whenever \( r \geq r_0(\varepsilon) \).

If we restrict ourselves to counting only so-called non-reciprocal, non-cyclotomic factors of \( F(x) \) then

\[
\sum_{\xi_i \text{ non-reciprocal}}^{I} m(i)[k(\xi_i) : Q(\xi_i)] \leq [k : Q] \frac{\log \mu(F)}{\log \theta_0}
\]

where \( \theta_0 = 1.324 \ldots \) is the Mahler measure of \( x^3 - x - 1 \).
Since often the Mahler measure $\mu(F)$ is not a practical parameter to use we note the chapter 2 inequalities

$$\log \mu(F) \leq \log H(F) \leq \log \nu(F) \approx \frac{\partial(F)}{r}.$$  

By non-reciprocal we mean here that $\xi_i$ and $1/\xi_i$ are not conjugate over $\mathbb{Q}$. Equivalently the minimal polynomial $g(x)$ for $\xi_i$ over $\mathbb{Q}$ is non-reciprocal in the usual sense (i.e. $x^{\delta(g)} g(1/x) \neq \pm g(x)$). The non-cyclotomic bounds and their equivalences will be discussed further in Section 3.2.

The rest of the results concern the number of cyclotomic factors. Theorems 2, 3 and 5 are perhaps the most interesting. The estimates of theorem 4 are straightforward and are given chiefly for comparison with theorems 2 and 5. The field dependent constants $c(k)$, $b(k)$ and $J$ are defined in (1.10), (1.11) and (1.12) above. A variety of bounds for $c(k)$ and $b(k)$ are discussed in section 4.2.

**Theorem 2** Let $F(x)$ be a polynomial factoring in $k[x]$ as indicated in (1.4).

Then for any $\varepsilon > 0$ the multiplicities of the cyclotomic factors of $F(x)$ satisfy

$$\sum_{n=1}^{\infty} \sum_{s=1}^{\delta(k;n)} e(n, s) \leq (1 + \varepsilon)\sqrt{c(k)} \frac{\partial(F)}{r} \left(\frac{\log r}{r}\right)^{1/2} \quad (1.20)$$

whenever $r \geq r_0(\varepsilon, k)$ is sufficiently large.

**Corollary 1** Let $F(x)$ be as above. Then for every $\varepsilon > 0$

$$\sum_{n=1}^{\infty} \frac{1}{\delta(k;n)} \sum_{s=1}^{\delta(k;n)} e(n, s) \leq (1 + \varepsilon) \left(\frac{\zeta(2)\zeta(3)}{\zeta(6)}\right)^{1/2} \frac{\partial(F)}{r} \left(\frac{\log r}{r}\right)^{1/2} \quad (1.21)$$

whenever $r > r_0(\varepsilon)$ is sufficiently large.
Theorem 3 Let \( F(x) \) be a polynomial in \( k[x] \) factoring in \( k[x] \) as above. Then we have the following bounds on the number of \( n \)-th and primitive \( n \)-th roots of unity amongst the roots of \( F(x) \):

For each positive integer \( n \leq r \)

\[
\sum_{m \mid n} \sum_{s=1}^{\delta(k;m)} e(m, s) \partial(\Phi_{m,s}) \ll \partial(F) \left( \frac{n}{r} \right)^{\frac{1}{2}} \tag{1.22}
\]

and for all \( n \) with \( \phi(n) \leq r \)

\[
\sum_{s=1}^{\delta(k;n)} e(n, s) \partial(\Phi_{n,s}) \ll \partial(F) \left( \frac{\phi(n)}{r} \right)^{\frac{1}{2}} \left( 1 + \left( \frac{\log \log 20n}{\log \left( \frac{r \log \log 20n}{\phi(n)} \right)} \right)^{\frac{1}{2}} \right) \tag{1.23}
\]

From these bounds we obtain immediately a restriction on the maximum multiplicity of a cyclotomic factor

\[
\max_{1 \leq n \leq \infty} \max_{1 \leq s \leq \delta(k;n)} e(n, s) \ll [k' : Q] \partial(F) r^{-\frac{1}{2}}. \tag{1.24}
\]

Theorem 4 Let \( F(x) \) be a polynomial in \( k[x] \) factoring as above. Counting without multiplicity the cyclotomic factors satisfy the following elementary bounds. For any \( \varepsilon > 0 \)

\[
\sum_{n=1}^{\infty} \sum_{s=1}^{\delta(k;n)} 1 \leq (1 + \varepsilon) \sqrt{2c(k)} \partial(F) \tag{1.25}
\]

\[
\sum_{n=1}^{\infty} 1 \leq (1 + \varepsilon) \sqrt{2b(k)} \partial(F) \tag{1.26}
\]

\[
\sum_{n=1}^{\infty} \frac{1}{\delta(k; n)} \sum_{s=1}^{\delta(k;n)} 1 \leq (1 + \varepsilon) \sqrt{2c(Q)} \partial(F) \tag{1.27}
\]
whenever $\partial(F) > N_0(\varepsilon, k)$ is sufficiently large. Note

$$c(Q) = b(Q) = \frac{\zeta(2)\zeta(3)}{\zeta(6)}.$$ 

**Theorem 5** Let $F(x) = \sum_{i=1}^{N(F)} a_i x^{n_i}$ be a polynomial in $k[x]$ with $N(F)$ non-zero coefficients. Suppose that $F(x)$ factors in $k[x]$ as above. Then, counting them without multiplicity, the cyclotomic factors of $F(x)$ satisfy

$$\sum_{\substack{n \geq 1 \\ e(n,s) \neq 0 \text{ for some } s}} 1 \ll d(J)\Delta(F)N(F). \quad (1.28)$$

Here $\Delta(F)$ is the complicated function defined in (1.7) with

$$\Delta(F) \ll \left( \sum_{i=1}^{N(F)} \sum_{j=1}^{N(F)} d(n_i - n_j) \right) \log N(F) \ll \partial(F)^c N(F)^2. \quad (1.29)$$

In addition the multiplicity of any factor of $F(x)$ (excluding possibly $x$) must satisfy

$$\max_{1 \leq i \leq l} m(j) < N(F)$$

and

$$\max_{1 \leq n \leq \infty} \max_{1 \leq s \leq \delta(k;n)} e(n,s) < N(F). \quad (1.30)$$

Using (1.3), (1.28) and (1.30) we obtain a whole set of alternative bounds for the quantities considered in (1.20), (1.21), (1.25) and (1.27).

$$\sum_{n=1}^{\infty} \frac{\delta(k;n)}{\delta(k;n)} \sum_{s=1}^{\delta(k;n)} e(n,s) \ll [k': Q]d(J)\Delta(F)N(F)^2$$

$$\sum_{n=1}^{\infty} \frac{1}{\delta(k;n)} \sum_{s=1}^{\delta(k;n)} e(n,s) \ll d(J)\Delta(F)N(F)^2$$
\[
\sum_{n=1}^{\infty} \frac{1}{\delta(k; n)} \sum_{c(n, s) \neq 0}^{\infty} 1 \ll \frac{[k' : Q]}{d(J) \Delta(F) N(F)}
\]

\[
\sum_{n=1}^{\infty} \frac{1}{\delta(k; n)} \sum_{c(n, s) \neq 0}^{\infty} 1 \ll d(J) \Delta(F) N(F').
\] (1.31)

### 1.3 A few brief comments on the theorems

Inequality (1.18) improves upon a result of Dobrowolski [9]; although (apart from the new field dependance) this improvement merely reflects a more efficient use of Dobrowolski's lower bound (3.1). It will be clear from the proof of theorem 1 that the existence of Lehmer's conjectured lower bound \(c > 1\) (such that all non-cyclotomic irreducible polynomials \(G(x) \neq x\) in \(\mathbb{Z}[x]\) have Mahler measure \(\log \mu(F) \geq \log c\) would lead at once to the bounds

\[
\sum_{i=1}^{I} m(i)[k(\xi_i) : Q(\xi_i)] \leq [k : Q] \frac{\log \mu(F)}{\log c}
\] (1.32)

\[
\leq [k : Q] \frac{\log H(F)}{\log c}
\] (1.33)

\[
\leq [k : Q] \frac{\log \nu(F)}{\log c}.
\] (1.34)

In fact we shall see in section 3.2 that conversely a bound of the form (1.33) or (1.34) would actually give us Lehmer's conjecture. Computer evidence suggests \(\mu(x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1) = 1.176 \ldots\) as the optimal \(c\). Indeed the inequality (1.19) (essentially Schinzel [21, theorem 1]) is a consequence of Smyth's proof of just such a lower bound for the subset of non-reciprocal polynomials in \(\mathbb{Z}[x]\). The polynomials

\[
F(x) = \prod_{i=1}^{I} (x^{3n_i} - x^{n_i} - 1)^{m(i)}
\]
show immediately the sharpness of (1.19). Section 3.2 will show the log $\theta_0$ in the secondary bounds

$$\sum_{i=1}^{I} m(i)[k(\xi_i) : Q(\xi_i)] \leq [k : Q] \frac{\log H(F)}{\log \theta_0} \quad (1.35)$$

$$\sum_{i=1}^{I} m(i)[k(\xi_i) : Q(\xi_i)] \leq [k : Q] \frac{\log \nu(F)}{\log \theta_0} \quad (1.36)$$

to be still the best possible constant – resolving a question posed by Schinzel [21, theorem 1]. There we demonstrate for each $m \in N$ the existence of polynomials $F_m(x)$ in $Q[x]$ with $(x^3 - x - 1)^m \mid F_m(x)$ having the comparatively low height

$$\log H(F_m) = m \log \theta_0 \left(1 + O\left(\frac{\log m}{m}\right)\right)$$

$$\log \nu(F_m) = m \log \theta_0 \left(1 + O\left(\frac{\log m}{m}\right)\right).$$

Now if $x^3 - x - 1$ factors in $k[x]$ as $x^3 - x - 1 = \prod_{j=1}^{d} g_j(x)$ these $F_m(x)$ plainly satisfy

$$\sum_{i=1}^{I} m(i)[k(\xi_i) : Q(\xi_i)] = [k : Q] \sum_{i=1}^{I} m(i) \frac{[k(\xi_i) : k]}{[Q(\xi_i) : Q]}$$

$$\geq [k : Q] \sum_{j=1}^{J} m \partial(g_j) = m[k : Q]$$

while our upper bounds (1.35) and (1.36) take the form

$$\leq [k : Q] m \left(1 + O\left(\frac{\log m}{m}\right)\right).$$

Since $m$ can be taken arbitrarily large the constant log $\theta_0$ must indeed have been optimal. Relying on a use of Siegel’s lemma the proof of the existence of these extremal $F_m(x)$ is unfortunately non-constructive.
As will be demonstrated by explicitly constructed examples in section 4.6 the bounds of theorem 2 are sharp – at least up to the determination of the precise constant. The field constant may indeed not be best possible. However we shall show in section 4.6 that we certainly cannot hope to decrease this upper bound by any absolute constant factor better than

\[
\frac{3}{\pi} = 0.954 \ldots
\]

for a general field. Even in the particular case \( k = \mathbb{Q} \) we cannot reduce it by anything smaller than

\[
\frac{3}{4} \left( \frac{3\zeta(6)}{\zeta(2)\zeta(3)} \right)^{\frac{1}{2}} = 0.931 \ldots
\]

When \( k = \mathbb{Q} \) both (1.23) and (1.24) coalesce; improving upon Schinzel’s version [22, theorem 2] by a factor of \( \log \log r \) and verifying his conjectured best possible bound. The theorem 2 inequalities will actually come from a more general weighted sum considered in section 4.3. However, apart from the two particular choices of weights giving rise to the theorem and its corollary, it is not obvious what others could be profitably used. As will be apparent from the proof we may drop the awkward condition \( r > r_0(\varepsilon, k) \) entirely; as long as we are willing to replace the ‘\( \leq \)’ by a ‘\( \ll \)’ and the limit defined constant \( c(k) \) by a corresponding supremum

\[
c(k)^* = \sup_{N \geq 1} \frac{1}{N} \sum_{\substack{n=1 \to \infty \\{k(\zeta_n) : k \}} \text{[} k(\zeta_n) : k \text{]}}.
\]

It is not hard to show that the upper bounds (1.16) for \( c(k) \) will still hold for \( c(k)^* \) if we again replace the ‘\( \leq \)’ with a ‘\( \ll \)’. More complicated variants of the
inequality incorporating an asymptotic 'error term' rather than an 'ε' or an explicit field dependence in the \( r_0(ε, k) \) are not difficult to concoct.

The estimates in theorem 3 would appear to be new; although when \( k = Q \) the corollary (1.24) was obtained by Schinzel [22, theorem 2]. Both results follow from much the more complicated bounds given in theorem 8. These more precise bounds are in fact best possible (up to a constant) as will be shown in section 4.6. However (although they mirror well the crucial relationship between the multiplicities \( e(n, s) \) and the \( e(np, s') \) where \( p \) is a prime) these formulae really do require too much detailed and usually inaccessible information to be generally applicable.

The bounds in theorem 4 should be regarded as the most basic that can be made in the non-multiplicity case. The first sum (1.25) counts the number of distinct irreducible cyclotomic factors of \( F(x) \) while (1.26) counts the number of \( n \) with \( F(ζ_n) = 0 \) for some primitive \( n \)-th root of unity \( ζ_n \). The remaining quantity (1.27) is simply a non-multiplicity version of the weighted, field independent sum (1.21). Clearly when \( k = Q \) all three quantities are identical. For the case \( k = Q \) a result of this order was given by Schinzel [21, theorem 3]. The polynomials

\[
G_{N,k}(x) = \prod_{\substack{n=1 \\ [k(ζ_n):k] \leq N}}^\infty \Phi_n(x)
\]

\[
H_{N,k}(x) = \prod_{\substack{n=1 \\ [k(ζ_n):k] \leq N}}^\infty \Phi_{n,1}(x)
\]

\[
L_N(x) = \prod_{\substack{n=1 \\ \phi(n) \leq N}}^\infty \Phi_n(x)
\]
explain the form and sharpness of the respective bounds.

Theorem 5 seems to be quite new. Indeed the optimal form of such a bound is not at all obvious. Since (as will be clear from theorem 10) we are really counting the number of cyclotomic factors of a whole set rather than of an individual polynomial, exactness seems implausible here. When \( k = \mathbb{Q} \) Schinzel [22] did in fact obtain via Mann's original theorem [18] that

\[
\sum_{n=1}^{\infty} 1 \ll \left( \frac{\vartheta(F) \log H(F)}{\log \log \vartheta(F)} \right)^{\frac{1}{2}}.
\]

In view of the relation \( N(F) \leq H(F)^2 \) of (2.24) and the \( \sqrt{\vartheta(F)} \) bounds of theorem 4 it is clear that even the roughest form of theorem 5

\[
\sum_{n=1}^{\infty} 1 \ll_{K, \varepsilon} \vartheta(F)^{\varepsilon} N(F)^3
\]

is a significant improvement upon such a bound. This is particularly so when \( N(F) \) is small. Indeed (1.37) supplants the theorem 4 bounds in the greatly extended range

\[
N(F) \ll_{K, \varepsilon} \vartheta(F)^{\frac{3}{2} - \varepsilon}.
\]

Similarly when counting with multiplicity the roughest theorem 5 bound

\[
\sum_{n=1}^{\infty} \sum_{s=1}^{\delta(kn)} e(n, s) \ll_{K, \varepsilon} \vartheta(F)^{\varepsilon} N(F)^4
\]

is not only better than the trivial bound \( \vartheta(F) \) when

\[
N(F) \ll_{\varepsilon, K} \vartheta(F)^{\frac{4}{3} - \varepsilon}
\]

but is in fact also always superior to our theorem 2 bound once

\[
N(F) \ll_{K, \varepsilon} \vartheta(F)^{\frac{5}{2} - \varepsilon}.
\]
Chapter 2

Basic definitions and inequalities

2.1 Preliminaries on valuations

Let $k$ be an algebraic number field. Then recall that a real-valued function $||$ on $k$ is called a valuation if for all $x$ and $y$ in $k$

(i) $|x| \geq 0$ with equality if and only if $x = 0$

(ii) $|xy| = |x||y|$

(iii) $|x + y| \leq |x| + |y|.$

When in addition the valuation satisfies the ultrametric or strong-triangle inequality:

(iii)' $|x + y| \leq \max \{|x|, |y|\}$

we say that $||$ is non-archimedean. Otherwise $||$ is called archimedean. For a non-archimedean valuation we have the following useful consequence of (iii)'

$$|x| < |y| \Rightarrow |x + y| = |y|. \quad (2.1)$$

For any field we always have the 'trivial valuation':

$$\|x\|_0 = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{else.} \end{cases}$$
When $k = \mathbb{Q}$ the usual absolute value gives an archimedean valuation on $\mathbb{Q}$. We shall denote this by $\| \|_\infty$. In addition, for each prime $p$, we obtain a non-archimedean valuation $\| \|_p$ on $\mathbb{Q}$ by

$$\| q \|_p = \begin{cases} p^{-1} & \text{if } q = p \\ 1 & \text{if } q \text{ is a prime, } q \neq p. \end{cases}$$

(Note: due to the total multiplicativity implied by condition (ii) it is sufficient here to simply specify behaviour at the primes).

Two valuations $| |_1$ and $| |_2$ are said to be equivalent if $| |_1 = | |_2^C$ for some real $C > 0$. It was shown early on by Ostrowski that every non-trivial valuation on $\mathbb{Q}$ must be equivalent either to one of the $\| \|_p$ or to $\| \|_\infty$. We let $V_Q$ denote the indexing set of places

$$V_Q = \{\infty, 2, 3, 5, \ldots\}$$

for these valuations on $\mathbb{Q}$. Here and throughout we shall use the ‘double-bar’ notation $\| \|_w$ for a valuation on a number field $k$ only if its restriction to $\mathbb{Q}$ coincides exactly with one of the $\| \|_v$, $v \in V_Q$. In general we let $V_k$ denote a set of places $w$ indexing a set of inequivalent valuations $\| \|_w$ on $k$ such that every non-trivial valuation on $k$ is equivalent to one of these.

If $M \subseteq k$ is a subfield of $k$ then restricting the valuation $\| \|_v$, $v \in V_k$ gives a valuation $\| \|_u$, $u \in V_M$ on $M$. The place $v$ is said to lie over the place $u$ and is denoted by $v | u$. In particular by restriction to $\mathbb{Q}$ either $v | \infty$ in which case $v$ is called infinite or archimedean or $v | p$ for some finite prime $p$ and $v$ is said to be finite or non-archimedean. The finite and infinite places $v$ will often be distinguished by the labels $v \not| \infty$ and $v | \infty$ respectively.
For a place \( v \in V_k \) we use \( k_v \) to denote the completion of \( k \) with respect to the metric

\[
d_v(x, y) = \| x - y \|_v.
\]

and \( \overline{k}_v \) to denote the algebraic closure of \( k_v \). It can be shown (see [7, chapter 7] or [14, chapter III]) that the absolute value \( v \) on \( k_v \) extends in a unique way to all the other elements \( \alpha \) of \( \overline{k}_v \) by taking

\[
\log \| \alpha \|_v = \frac{1}{[k_v(\alpha) : k_v]} \log \| \text{Norm}_{k_v(\alpha)/k_v}(\alpha) \|_v.
\]

We let \( \Omega_v \) denote the completion of \( \overline{k}_v \) with respect to this valuation. \( \Omega_v \) is then both complete and algebraically closed (see for example [14, chapter III]).

Now let \( L = k(\alpha) \) be an extension of \( k \). Then each valuation \( \| \cdot \|_v \), \( v \in V_k \) can be extended to a set of valuations \( \| \cdot \|_w \) on \( L \). Construction of these \( w \in V_L \), \( w | v \) is quite concrete. Let \( K_v = k_v(\alpha_1, \alpha_2, \ldots, \alpha_n) \) where \( \alpha_1, \ldots, \alpha_n \) are the \( n = [L : k] \) conjugates of \( \alpha = \alpha_1 \) over \( k \) and let \( M_L/k \) denote the \( k \)-monomorphisms \( \sigma_i : L \to \overline{Q} \) given by \( \sigma_i : \alpha \mapsto \alpha_i \). Now as observed above

\[
\log \| \beta \|_{w_0} = \frac{1}{[k_v : k_v]} \log \| \text{Norm}_{K_v/k_v}(\beta) \|_v
\]

produces a valuation on \( K_v \) extending \( \| \cdot \|_v \) and moreover such a construction is the only way to extend \( \| \cdot \|_v \) from the complete field \( k_v \) to \( K_v \). However each \( \sigma_i \in M_L/k \) gives us a new way of embedding \( L \) in \( K_v \) and hence each produces a valuation

\[
\log \| \beta \|_{w_{\sigma_i}} = \log \| \sigma_i(\beta) \|_{w_0} = \frac{1}{[k_v(\alpha_i) : k_v]} \log \| \text{Norm}_{k_v(\alpha_i)/k_v}(\sigma_i(\beta)) \|_v
\]

for all \( \beta \in L \). It is not hard to show that all extensions must come from one or other of these embeddings (see for example [7, chapter 9]). Of course in general
all these $||\ +w_r$ may not be distinct. In fact it is readily seen that

$$L_{w_r} \cong k_v(\tau(\alpha))$$

with

$$||\ +w_r = ||\ +w_r$$

if and only if $\tau(\alpha)$ is one of the $[k_v(\sigma(\alpha)) : k_v] = [L_{w_u} : k_{w_r}]$ conjugates of $\sigma(\alpha)$ over $k_v$.

In summary, if the minimal polynomial $\phi(x)$ for $\alpha$ over $k$ factors into irreducibles as $\phi(x) = \phi_1(x) \cdots \phi_M(x)$ over $k_v[x]$ this construction produces $M$ distinct extensions $||\ +w_m$ of $||\ +v$ to $L = k(\alpha)$ with multiplicity $\partial(\phi_m) = [L_{w_m} : k_{w_m}]$

respectively. In particular we see at once that

$$\sum_{\substack{w|v \\
w \in V_L}} [L_w : k_w] = [L : k]. \quad (2.2)$$

In order to work only with distinct valuations and yet not to lose these multiplicities we define a normalised ‘single-bar’ valuation:

$$|\beta|_w = ||\ +w^{d_w/d} \quad (2.3)$$

where

$$d_w = [L_w : Q_w] \text{ and } d = [L : Q]. \quad (2.4)$$

Note: in view of (2.2) we have

$$\sum_{\substack{w|v \\
w \in V_L}} \frac{d_w}{d} = 1 \quad (2.5)$$
for all $v \in V_Q$.

An immediate advantage of these normalised valuations is the following product formula:

**Lemma 1 (The Product Formula)** Suppose $\alpha \in k \setminus \{0\}$. Then

$$\sum_{w \in V_k} \log |\alpha|_w = 0. \quad (2.6)$$

**Proof:** The product formula on $Q$ is familiar and straightforward:

If

$$n = \pm p_1^{\alpha_1} \cdots p_r^{\alpha_r}$$

where the $p_i$ are distinct primes and the $\alpha_i$ non-zero integers, then

$$\sum_{v \in V_Q} \log |n|_v = \log \|n\|_\infty + \sum_{i=1}^r \log \|n\|_{p_i}$$

$$= \log (p_1^{\alpha_1} \cdots p_r^{\alpha_r}) + \sum_{i=1}^r \log p_i^{-\alpha_i} = 0.$$

For a general $k$

$$\sum_{w \in V_k} \log |\alpha|_w = \sum_{v \in V_Q} \sum_{w|v} \frac{[k_w : Q_w]}{[k : Q]} \log \|\alpha\|_w$$

$$= \frac{1}{[k : Q]} \sum_{v \in V_Q} \sum_{\sigma \in \mathcal{M}_k/Q} \log \|\alpha\|_{w,\sigma}$$

$$= \frac{1}{[k : Q]} \sum_{v \in V_Q} \log \|\text{Norm}_{k/Q} \alpha\|_v$$

and the result follows from the rational case.
Given a \( v \in V_k \) and an extension \( \mathbb{L} \supseteq k \) the valuations \( w \mid v \) in \( \mathbb{V}_L \) plainly satisfy

\[
|\alpha|_w = \|\alpha\|_v \frac{[\mathbb{L}_w : k_w]}{[\mathbb{L} : k]} = \frac{[\mathbb{L}_w : k_w]}{[\mathbb{L} : k]} |\alpha|_v
\]  

(2.7)

for all \( \alpha \in k \). As a consequence of (2.2) and (2.7) we observe the additional useful property of the single-bar valuations:

For any \( \alpha \in k \subseteq \mathbb{L} \) and place \( v \in V_k \)

\[
\sum_{w \mid v, w \in \mathbb{V}_L} \log |\alpha|_w = \left( \sum_{w \mid v, w \in \mathbb{V}_k} \frac{[\mathbb{L}_w : k_w]}{[\mathbb{L} : k]} \right) \log |\alpha|_v = \log |\alpha|_v.
\]  

(2.8)

Similarly

\[
\sum_{w \mid v, w \in \mathbb{V}_L} \log^+ |\alpha|_w = \log^+ |\alpha|_v
\]  

(2.9)

where \( \log^+ x = \max\{0, \log x\} \).

These properties will become important later on in removing any ambiguity over field dependance from our definitions.

Now from our construction of the absolute values we easily obtain the following lemma:

**Lemma 2** Given an extension \( k \subseteq \mathbb{L} \) and a \( k \)-monomorphism \( \tau : \mathbb{L} \rightarrow \overline{Q} \).

Then

\[
\sum_{w \mid v, w \in \mathbb{V}_L} \log |\tau(\alpha)|_w = \sum_{w \mid v, w \in \mathbb{V}_L} \log |\alpha|_w
\]  

(2.10)

for all places \( v \in V_k \) and all \( \alpha \in \mathbb{L} \setminus \{0\} \) with \( \tau(\alpha) \in \mathbb{L} \).

A similar relation holds with the log replaced by \( \log^+ \).
Proof of lemma 2: Since $\tau$ merely permutes the $k$-monomorphisms we see that

$$
\sum_{\substack{w|v \\ w \in V_L}} \log |\tau(\alpha)|_w = \sum_{\substack{w|v \\ w \in V_L}} \frac{[L_w : k_w]}{[L : k]} \log \|\tau(\alpha)\|_w
$$

$$
= \frac{1}{[L : k]} \sum_{\sigma \in M_{L/k}} \log \|\tau(\alpha)\|_{w_{\sigma}}
$$

$$
= \frac{1}{[L : k]} \sum_{\sigma \in M_{L/k}} \log \|\sigma \tau(\alpha)\|_{w_0}
$$

$$
= \frac{1}{[L : k]} \sum_{\sigma \in M_{L/k}} \log \|\sigma(\alpha)\|_{w_0}
$$

$$
= \sum_{\substack{w|v \\ w \in V_L}} \log |\alpha|_w
$$

as required. Similarly for the $\log^+$. 

Finally we note that for any $\alpha \in k$ we have $\|\alpha\|_w = 1$ for all but finitely many $w \in V_k$. This is straightforward when $\alpha$ is in $Q$ (clearly there are only finitely many $v \in V_Q$ with $\|\alpha\|_v \neq 1$ and hence only finitely many extensions $w \in V_k$, $w | v$ of these to $k$). Therefore if $\sum_{i=0}^{L} a_i x^i$ is the minimal polynomial for $\alpha$ over $Q$

$$
S = \{w \in V_k : w | \infty \text{ or } \|a_i\|_w \neq 1 \text{ for some } 0 \leq i \leq L\}
$$

is a finite set. Now, by the ultrametric inequality, any $w \in V_k \setminus S$ with $\|\alpha\|_w \neq 1$ would have to satisfy

$$
0 = \| \sum_{i=1}^{L} a_i x^i \|_w = \begin{cases} 
\|\alpha\|_w^L & \text{if } \|\alpha\|_w > 1 \\
1 & \text{if } \|\alpha\|_w < 1
\end{cases}
$$

— which plainly cannot occur. Hence $\|\alpha\|_w = 1$ for all $w \in V_k$ not in $S$. 
Armed with these valuations we are now ready to define our various heights and measures on polynomials.

2.2 Heights and measures on polynomials

Let $v$ be a place of $k$ and let

$$F(x) = \sum_{i=0}^{L} a_i x^i = a_L \prod_{i=1}^{L} (x - \alpha_i) \quad (2.11)$$

be a non-trivial polynomial in $\Omega_v[x]$. Then we define the local heights $H_v$ and $\nu_v$ on $F(x)$ by

$$H_v(F) = \begin{cases} \max_{0 \leq i \leq L} |a_i|_v & \text{if } v \nmid \infty \\
\left( \sum_{i=0}^{L} ||a_i||_v^2 \right)^{d_v/2d} & \text{if } v | \infty \end{cases} \quad (2.12)$$

and

$$\nu_v(F) = \sup \{|F(z)|_v : z \in \Omega_v, |z|_v = 1\}. \quad (2.13)$$

We define the local Mahler measure $\mu_v$ by

$$\log \mu_v(F) = \log |a_L|_v + \sum_{i=1}^{L} \log^+ |a_i|_v. \quad (2.14)$$

Now suppose that $F(x)$ is in $k[x]$ then plainly $H_v(F), \nu_v(F)$ and $\mu_v(F)$ are all defined for each place $v$ of $k$. In addition it is not hard to see that each of them must be 1 for all but finitely many of the $v$. Hence we can define the global heights $H, \nu$ and global Mahler measure $\mu$ by

$$H(F) = \prod_{v \in \mathcal{V}_k} H_v(F) \quad (2.15)$$

$$\nu(F) = \prod_{v \in \mathcal{V}_k} \nu_v(F) \quad (2.16)$$
and

$$\mu(F) = \prod_{u \in V_k} \mu_u(F). \quad (2.17)$$

We first need to check that these definitions are independent of the particular field $k$ chosen. Clearly it is enough to look at a field $L \supseteq k$. As a consequence of our normalisations (2.3) we see that

$$H_w(F) = H_v(F)^{[L_w:k_w]/[L:k]}$$
$$\nu_w(F) = \nu_v(F)^{[L_w:k_w]/[L:k]}$$
$$\mu_w(F) = \mu_v(F)^{[L_w:k_w]/[L:k]}$$

for all $v \in V_k$ and $w \in V_L$ with $w \mid v$. Thus, as in (2.8), we see from (2.2) that

$$\prod_{w \mid v \atop w \in V_L} H_w(F) = H_v(F)$$
$$\prod_{w \mid v \atop w \in V_L} \nu_w(F) = \nu_v(F)$$
$$\prod_{w \mid v \atop w \in V_L} \mu_w(F) = \mu_v(F)$$

and the desired invariance is plain.

Hence these heights and measures may be genuinely regarded as functions on the whole of $\overline{Q}[x] \setminus \{0\}$. For completeness we assume that $H(G) = \nu(G) = \mu(G) = 0$ for the identically zero polynomial. Clearly for any non-zero $\gamma$ in $k$ we have

$$H_v(\gamma F) = |\gamma|_v H_v(F)$$
$$\nu_v(\gamma F) = |\gamma|_v \nu_v(F)$$
\[ \mu_v(\gamma F) = |\gamma|_v \mu_v(F). \]

In particular the product formula (2.6) gives us at once the useful homogeneity of these functions:

\[ \begin{align*}
H(\gamma F) &= H(F) \\
\nu(\gamma F) &= \nu(F) \\
\mu(\gamma F) &= \mu(F).
\end{align*} \tag{2.18} \]

for all \( \gamma \) in \( \overline{\mathbb{Q}} \setminus \{0\} \).

It is of course natural to ask how these various quantities are related. Many of these results can be found in Bombieri [3]. The non-archimedean case is particularly simple:

**Lemma 3** Let \( F(x) \) be a polynomial in \( \Omega_v[x] \) where \( v \nmid \infty \). Then

\[ \mu_v(F) = \nu_v(F) = H_v(F). \tag{2.19} \]

**Proof of lemma 3:** Let

\[ \mathcal{R}_v = \{ z \in \Omega_v : |z|_v \leq 1 \} \]

\[ \mathcal{M}_v = \{ z \in \Omega_v : |z|_v < 1 \} \]

then \( \mathcal{M}_v \) is of infinite index in \( \mathcal{R}_v \). This it not difficult to see; for example if \( q \) is a prime with \( v \nmid q \) then by considering the discriminant of \( \Phi_{q^n}(x) \) it is clear that all the \( \phi(q^n) \) primitive \( q^n \)-th roots of unity must lie in different cosets – and \( n \) can be chosen as large as we like.
Suppose
\[ F(x) = a_L \prod_{i=1}^{L} (x - \alpha_i) = \sum_{i=0}^{L} a_i x^i \]
then by the above comments we are assured of a \( z_0 \in \Omega_v \) with \( |z_0|_v = 1 \) and \( |z_0 - \alpha_i|_v = 1 \) for all the roots \( \alpha_i \) in \( \mathcal{R}_v \). Hence from (2.1)

\[ |z_0 - \alpha_i|_v = \max \{ 1, |\alpha_i|_v \} \]

for all the \( \alpha_i \) and

\[
\mu_v(F) = |a_L|_v \prod_{i=1}^{L} \max \{ 1, |\alpha_i|_v \}
\]

\[ = |a_L|_v \prod_{i=1}^{L} |z_0 - \alpha_i|_v \]

\[ = |F(z_0)|_v \]

\[ \leq \sup_{|z|_v = 1} |F(z)|_v = \nu_v(F). \]

Now
\[ \nu_v(F) = \sup_{|z|_v = 1} |\sum_{i=0}^{L} a_i x^i|_v \]

\[ \leq \max_{0 \leq i \leq L} |a_i|_v = H_v(F) \]

while
\[ H_v(F) = |a_{L-m}|_v = |a_L \sum_{i_1, \ldots, i_m \subseteq \{1, \ldots, L\}} \alpha_{i_1} \cdots \alpha_{i_m}|_v \]

\[ \leq |a_L|_v \max |\alpha_{i_1} \cdots \alpha_{i_m}|_v \]

\[ \leq |a_L|_v \prod_{i=1}^{L} \max \{ 1, |\alpha_i|_v \} = \mu_v(F). \]

So
\[ \mu_v(F) = \nu_v(F) = H_v(F) \]
as claimed. The equality $H_v(F) = \mu_v(F)$ is often referred to as Gauss' lemma.

For the infinite places we gain the following string of inequalities:

**Lemma 4** Let $F(x)$ be a polynomial in $\Omega_v[x]$ where $v | \infty$. Then

$$H_v(F)2^{-\beta(F)d_v/d} \leq \mu_v(F) \leq H_v(F)$$

$$\leq \nu_v(F) \leq H_v(F)N(F)^{d_v/2d}. \quad (2.20)$$

**Proof of lemma 4:** When $v | \infty$ we let $\mathcal{U}_v$ denote the compact group of units in $\Omega_v$

$$\mathcal{U}_v = \{z \in \Omega_v : |z|_v = 1\}$$

and let $m_v$ denote Haar measure on $\mathcal{U}_v$ normalised so that $m_v(\mathcal{U}_v) = 1$. Now by Jensen’s formula:

$$\log^+ |\alpha|_v = \int_{\mathcal{U}_v} \log |z - \alpha|_v dm_v(z)$$

giving us the alternative formulation of $\mu_v(F)$

$$\log \mu_v(F) = \int_{\mathcal{U}_v} \log |F(z)|_v dm_v(z) \quad (2.21)$$

taking $v | \infty$.

Suppose that

$$F(x) = \sum_{i=0}^{L} a_i x^i = a_L \prod_{i=1}^{L} (x - \alpha_i).$$

Then from the definitions of $H_v$ and $\mu_v$:

$$H_v(F)^{2d/d_v} = \|a_L\|_v^2 \sum_{m=0}^{L} \| \sum_{1 \leq i_1 < \cdots < i_m \leq L} \alpha_{i_1} \cdots \alpha_{i_m} \|_v^2$$
\[
\begin{aligned}
\leq & \ \|a_L\|^2 \prod_{l=0}^{L} \max \left\{ 1, \|\alpha_l\|^2 \right\} \sum_{m=0}^{L} \left( \sum_{0 \leq i_1 < \cdots < i_m} 1 \right)^2 \\
= & \ \mu_v(F)^{2d/d_v} \sum_{m=0}^{L} \binom{L}{m}^2 \\
< & \ \mu_v(F)^{2d/d_v} 4^L
\end{aligned}
\]

and the first inequality follows. From (2.21), the concavity of the logarithm (Jensen’s Inequality), and Parseval’s formula we obtain the second inequality:

\[
\begin{aligned}
\log \mu_v(F) &= \int_{\mathcal{U}_v} \log |F(z)|_v dm_v(z) \\
&= \frac{d_v}{2d} \int_{\mathcal{U}_v} \log \|F(z)\|_v^2 dm_v(z) \\
&\leq \frac{d_v}{2d} \log \int_{\mathcal{U}_v} \|F(z)\|_v^2 dm_v(z) \\
&= \frac{d_v}{2d} \log \left( \sum_{i=0}^{L} \|a_i\|_v^2 \right) = \log H_v(F).
\end{aligned}
\]

Now from Parseval’s formula and the Cauchy-Schwartz inequality we see that

\[
\begin{aligned}
H_v(F)^{2d/d_v} &= \sum_{i=0}^{L} \|a_i\|_v^2 = \int_{\mathcal{U}_v} \|F(z)\|_v^2 dm_v(z) \\
&\leq \sup_{z \in \mathcal{U}_v} \|F(z)\|_v^2 = \nu_v(F)^{2d/d_v} \\
&\leq \left( \sum_{i=0}^{L} \|a_i\|_v^2 \right) \left( \sum_{i=0}^{L} 1 \right) \\
&= H_v(F)^{2d/d_v} N(F)
\end{aligned}
\]

and the remaining inequalities are clear.

We also have the corresponding global inequalities.
Lemma 5 Let $F(x)$ be a polynomial in $k[x]$. Then

$$H(F)2^{-9(F)} \leq \mu(F) \leq H(F) \leq \nu(F) \leq H(F)N(F)^{\frac{1}{2}} \leq H(F)^2. \tag{2.22}$$

Proof of lemma 5: The first four inequalities follow from simply multiplying the local inequalities (2.19) and (2.20) over all the places $v \in V_k$, since recall from (2.5)

$$\sum_{v | \infty \atop v \in V_k} \frac{d_v}{d} = 1. \tag{2.23}$$

The last inequality will follow from the bound:

$$N(F) \leq H(F)^2. \tag{2.24}$$

When $F(x)$ is in $\mathbb{Z}[x]$ such an estimate is obvious. To see that it still holds in an arbitrary number field we first apply the product formula (2.6)

$$N(F) = \sum_{i=0}^{L} \left( \prod_{v \in V_k} |a_i|_{v}^2 \right)$$

$$\leq \left( \prod_{v | \infty \atop v \in V_k} \max_{0 \leq i \leq L} |a_i|_{v}^2 \right) \left( \sum_{i=0}^{L} \prod_{v | \infty \atop v \in V_k} |a_i|_{v}^2 \right)$$

$$= \left( \prod_{v | \infty \atop v \in V_k} H_v(F)^2 \right) \left( \sum_{i=0}^{L} \prod_{v | \infty \atop v \in V_k} (|a_i|_{v}^2)^{d_v/d} \right).$$

Then, in view of (2.23), we can use Hölder’s inequality on the inner sum

$$\sum_{i=0}^{L} \prod_{v | \infty \atop v \in V_k} (|a_i|_{v}^2)^{d_v/d} \leq \prod_{v | \infty \atop v \in V_k} \left( \sum_{i=0}^{L} (|a_i|_{v}^2)^{d_v/d} \right)$$

$$= \prod_{v | \infty \atop v \in V_k} H_v(F)^2$$
and (2.24) is clear.

We note two elementary yet essential properties of \( \mu \):

(i) \( \log \mu(FG) = \log \mu(F) + \log \mu(G) \)

(ii) \( \log \mu(F) \geq 0 \)

for all \( F(x) \) and \( G(x) \) in \( \overline{\mathbb{Q}}[x] \setminus \{0\} \). Both of these are immediate from the definition of \( \mu \) (since by the product formula \( \sum_v \log |a_L|_v = 0 \)). From inequality (2.22) the non-negativity is plainly shared by \( \log \nu \) and \( \log H \). The additivity though is not; the behaviour of \( H \) being particular complicated with respect to multiplication. With the \( \nu \) height we can at least say that

\[
\log \nu(F_1F_2) \leq \log \nu(F_1) + \log \nu(F_2)
\]

with

\[
\log \nu(F^m) = m \log \nu(F).
\]

This partly explains the preference sometimes shown here for the \( \nu \) over the \( H \) height. Of course \( \mu \) suffers from being the least tangible of the three.

It remains to check as claimed in the introduction that these heights do coincide with the more familiar definitions when \( k = \mathbb{Q} \). In view of (2.8) we may assume after multiplication by an appropriate constant that

\[
F(x) = \sum_{i=0}^{L} a_i x^i = a_L \prod_{i=1}^{L} (x - \alpha_i)
\]

with the \( a_i \in \mathbb{Z} \) and

\[
gcd(a_0, a_1, \ldots, a_L) = 1.
\]
From the coprimeness we see immediately that $H_p(F) = 1$ at all the finite places and hence from (2.9) that

$$H_p(F) = \nu_p(F) = \mu_p(F) = 1$$

for all primes $p$. So everything depends solely on the infinite place and

$$H(F) = H_\infty(F) = \left( \sum_{i=0}^{L} \|a_i\|_\infty^2 \right)^{\frac{1}{2}}$$

$$\nu(F) = \nu_\infty(F) = \sup_{\|x\|_\infty = 1} \|F(x)\|_\infty$$

and

$$\mu(F) = \mu_\infty(F) = \|a_L\|_\infty \prod_{i=1}^{L} \max \{1, \|\alpha_i\|_\infty\}$$

as expected.

By construction $H, \nu$ and $\mu$ are also invariant under elements of $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$:

**Lemma 6** Let $\tau$ be an element of $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and $F(x)$ a polynomial in $\overline{\mathbb{Q}}[x]$. Then

$$\mu(F) = \mu(\tau F)$$

$$H(F) = H(\tau F)$$

$$\nu(F) = \nu(\tau F).$$

**Proof of lemma 6:** Let $L$ be a splitting field for $F(x)$. Then from lemma 2 and the definition (2.4) for $\log \mu_v$

$$\sum_{\substack{w|v, \mathfrak{m} w \in \mathfrak{v}_L \mathfrak{v}_L}} \log \mu_w(F) = \sum_{\substack{w|v, \mathfrak{m} w \in \mathfrak{v}_L \mathfrak{v}_L}} \log \mu_w(\tau F)$$
for each place $u$ of $Q$. Summing over all the places of $L$ gives the first equality.

As $\nu_w$ and $H_w$ coincide with $\mu_w$ at all the finite places we need only check that a similar relationship holds for the infinite places of $L$. Since $\tau$ merely permutes the elements of $M_{L/Q}$ we have

$$
\sum_{w|\infty \atop w \in V_L} \log H_w(F) = \sum_{w|\infty \atop w \in V_L} \left[ \frac{[L_w : Q_w]}{2[L : Q]} \right] \log \left( \sum_{i=0}^{L} \|a_i\|_w^2 \right)
$$

$$
= \frac{1}{2[L : Q]} \sum_{\sigma \in M_{L/Q}} \log \left( \sum_{i=0}^{L} \|\sigma(a_i)\|_\infty^2 \right)
$$

$$
= \frac{1}{2[L : Q]} \sum_{\sigma \in M_{L/Q}} \log \left( \sum_{i=0}^{L} \|\sigma \tau(a_i)\|_\infty^2 \right)
$$

$$
= \sum_{w|\infty \atop w \in V_L} \log H_w(\tau F)
$$

and similarly

$$
\sum_{w|\infty \atop w \in V_L} \log \nu_w(F) = \sum_{w|\infty \atop w \in V_L} \left[ \frac{[L_w : Q_w]}{[L : Q]} \right] \log \left( \sup_{z \in \mathcal{U}_w} \||F(z)||_w \right)
$$

$$
= \frac{1}{[L : Q]} \sum_{\sigma \in M_{L/Q}} \log \left( \sup_{z \in \mathcal{C}} \||\sigma F(z)||_\infty \right)
$$

$$
= \frac{1}{[L : Q]} \sum_{\sigma \in M_{L/Q}} \log \left( \sup_{z \in \mathcal{C}} \||\sigma \tau F(z)||_\infty \right)
$$

$$
= \sum_{w|\infty \atop w \in V_L} \log \nu_w(\tau F).
$$

The lemma follows at once on summing over all the places of $Q$.

Finally in this section we give some inequalities relating the height of a polynomial to the height of its derivatives. We shall state them in terms of
the $\nu$ height since this turns out to be the most convenient to use later on. We write $D^n$ for the $n$-fold differential operator:

$$D^n = \frac{1}{n!} \left( \frac{d}{dx} \right)^n.$$ 

Lemma 7 Let $F(x) = \sum_{i=0}^{L} a_i x^i$ be a polynomial in $\Omega_v[x]$. Then for each positive integer $n$

$$\nu_v(D^n F) \leq \begin{cases} \nu_v(F) & \text{if } v \not\| \infty \\ \left( \frac{\partial(F)}{n} \right)^{d_v/d} \nu_v(F) & \text{if } v \| \infty. \end{cases}$$ (2.25)

Proof of Lemma 7: When $n > \partial(F) + 1$ the bounds are vacuously true, so we shall assume that $n \leq \partial(F) + 1$. Clearly then

$$D^n F(x) = \sum_{i=n}^{L} \binom{i}{n} a_i x^{i-n}.$$ 

Hence when $v \not\| \infty$ we see from Lemma 3 that

$$\nu_v(D^n F) = H_v(D^n F) = \max_{n \leq i \leq L} \left| \binom{i}{n} a_i \right|_v$$

$$\leq \max_{n \leq i \leq L} |a_i|_v \leq H_v(F) = \nu_v(F).$$

If $v \| \infty$ then by a theorem of Bernstein [26, p11]

$$\nu_v(D^1 F) = \left( \sup_{\|x\|_v = 1} \|F'(z)\|_v \right)^{d_v/d}$$

$$\leq \left( \frac{\partial(F)}{n} \sup_{\|z\|_v = 1} \|F(z)\|_v \right)^{d_v/d}$$

$$= \partial(F)^{d_v/d} \nu_v(F).$$

The general result follows by an $n$-fold application of the above.
Lemma 8 Let $F(x)$ be a polynomial in $\Omega_v[x]$, $n$ a positive integer and $\xi$ an element of $\Omega_v$. Then

$$|D^n F(\xi)|_v \leq \begin{cases} \nu_v(F) \max \{1, |\xi|_v\}^{\frac{\partial(F)}{d} - n} & \text{if } v \not\mid \infty \\ \nu_v(F) \left(\frac{\partial(F)}{n}\right)^{d_v/d} \max \{1, |\xi|_v\}^{\frac{\partial(F)}{n} - n} & \text{if } v \mid \infty. \end{cases} \quad (2.26)$$

Proof of lemma 8: In view of lemma 7 it plainly suffices to prove the result with $n = 0$. When $v \not\mid \infty$ the ultrametric inequality gives

$$|F(\xi)|_v \leq \max_{0 \leq l \leq L} |a_i \xi^l|_v \leq \left(\max_{0 \leq l \leq L} |a_l|_v\right) \left(\max_{0 \leq l \leq L} |\xi|^l_v\right)$$

$$= H_v(F) \max \{1, |\xi|_v\}^L = \nu_v(F) \max \{1, |\xi|_v\}^L.$$

When $v \mid \infty$ we appeal the maximum modulus theorem:

If $|\xi|_v \leq 1$ we obtain

$$|F(\xi)|_v \leq \sup_{|z|_v \leq 1} |F(z)|_v = \sup_{|z|_v = 1} |F(z)|_v = \nu_v(F).$$

Similarly if $|\xi|_v > 1$ writing $G(x)$ for the reciprocal polynomial

$$G(x) = x^{\partial(F)} F(1/x)$$

we obtain

$$|F(\xi)|_v = |\xi|^L_v |G(1/\xi)|_v \leq |\xi|^L_v \sup_{|z|_v \leq 1} |G(z)|_v$$

$$= |\xi|^L_v \sup_{|z|_v = 1} |G(z)|_v = |\xi|^L_v \sup_{|z|_v = 1} |F(z)|_v = |\xi|^L_v \nu_v(F).$$

This completes the proof of the lemma.

In the next section we define the $\mathcal{L}$ operators so crucial to the proof of theorem 2.
2.3 The operator $\mathcal{L}$ and some of its properties

We suppose that $\nu$ is a place of $k$ with $\nu \nmid \infty$. Given two non-trivial polynomials in $\Omega_\nu[x]$

\[
F(x) = a_L \prod_{l=1}^{L} (x - \alpha_l)
\]
\[
G(x) = b_M \prod_{m=1}^{M} (x - \beta_m)
\]

we define the local operator $\mathcal{L}_\nu$ by

\[
\mathcal{L}_\nu(F, G) = \sum_{l=1}^{L} \sum_{m=1}^{M} \left\{ \log^+ |\alpha_l|_\nu + \log^+ |\beta_m|_\nu - \log |\alpha_l - \beta_m|_\nu \right\} .
\]

(2.27)

$\mathcal{L}_\nu$ has the following elementary properties

(i) $\mathcal{L}_\nu(F, G) = \mathcal{L}_\nu(G, F)$

(ii) $\mathcal{L}_\nu(\gamma F, G) = \mathcal{L}_\nu(F, G)$ for all $\gamma \neq 0$ in $\Omega_\nu$

(iii) $\mathcal{L}_\nu(F_1 F_2, G) = \mathcal{L}_\nu(F_1, G) + \mathcal{L}_\nu(F_2, G)$

(iv) $\mathcal{L}_\nu(F, G) \geq 0$.

The symmetry, homogeneity and additivity are immediate from the definition. The non-negativity follows from the ultrametric inequality:

\[
\log |\alpha_l - \beta_m|_\nu \leq \log (\max \{|\alpha_l|_\nu, |\beta_m|_\nu\}) \leq \log^+ |\alpha_l|_\nu + \log^+ |\beta_m|_\nu.
\]

Lemma 9 Let $\nu$ be a place of $k$ with $\nu \nmid \infty$. Suppose $F(x)$ and $G(x)$ are non-zero polynomials in $\Omega_\nu[x]$ with no common factors. Then $\mathcal{L}_\nu$ takes the simpler form

\[
\mathcal{L}_\nu(F, G) = \partial(G) \log H_\nu(F) + \partial(F) \log H_\nu(G) - \log |\text{Res}(F, G)|_\nu
\]

(2.29)
where \( \text{Res}(F, G) \) denotes the resultant of \( F(x) \) and \( G(x) \). If \( F(x) \) has no multiple zeros then

\[
\mathcal{L}_v(F, F) = 2(\partial(F) - 1) \log H_v(F) - \log |\text{Disc}(F)|_v
\]  

(2.30)

where \( \text{Disc}(F) \) denotes the discriminant of \( F(x) \).

**Proof of lemma 9:** If \( F(x) \) and \( G(x) \) are as in \( (2.27) \) then by \( (2.14) \) and lemma 3

\[
\mathcal{L}_v(F, G) = \partial(G) \sum_{i=1}^L \log^+ |\alpha_i|_v + \partial(F) \sum_{m=1}^M \log^+ |\beta_m|_v - \sum_{i=1}^L \sum_{m=1}^M \log |\alpha_i - \beta_m|_v
\]

\[
= \partial(G) (\log \mu_v(F) - \log |a_L|_v) + \partial(F) (\log \mu_v(G) - \log |b_M|_v)
\]

\[
- \log |\prod_{i=1}^L \prod_{m=1}^M (\alpha_i - \beta_m)|_v
\]

\[
= \partial(G) \log H_v(F) + \partial(F) \log H_v(G) - \log |a_L^{\partial(G)} b_M^{\partial(F)} \prod_{i=1}^L \prod_{m=1}^M (\alpha_i - \beta_m)|_v
\]

and \( (2.29) \) follows from the definition of the resultant.

Similarly

\[
\mathcal{L}_v(F, F) = 2(\partial(F) - 1) \sum_{i=1}^L \log^+ |\alpha_i|_v - \sum_{i=1}^L \sum_{j=1}^L \log |\alpha_i - \alpha_j|_v
\]

\[
= 2(\partial(F') - 1) \log H_v(F') - \log |a_L^{2(\partial(F') - 1)} \prod_{i=1}^L \prod_{i<j}^L (\alpha_i - \alpha_j)^2|_v
\]

and \( (2.30) \) follows from the definition of the discriminant.

Now suppose \( F(x) \) and \( G(x) \) are non-zero polynomials splitting completely in \( k \). Then \( \mathcal{L}_v(F, G) \) is defined at all the finite places of \( k \). Moreover it is plain from lemma 6 that any two irreducible \( F(x) \) and \( G(x) \) satisfy
\[ \mathcal{L}_v(F, G) = 0 \] for all but finitely many places \( v \). From the additive property (iii) the same must plainly be true for an arbitrary pair of non-zero polynomials. Hence we can define a global function \( \mathcal{L} : (k[x] \setminus \{0\}) \times (k[x] \setminus \{0\}) \to [0, \infty) \) by

\[ \mathcal{L}(F, G) = \sum_{v \in V_k, v \nmid \infty} \mathcal{L}_v(F, G). \tag{2.31} \]

As before we should check that such a definition is independent of the choice of field \( k \). Suppose that \( k \subseteq L \). Then from (2.8), (2.9) and the definition of \( \mathcal{L}_w \) we see immediately that for all \( v \in V_k, v \nmid \infty \)

\[ \sum_{w \in V_L, w \nmid v} \mathcal{L}_w(F, G) = \mathcal{L}_v(F, G). \]

The desired field invariance follows at once. In particular we may regard \( \mathcal{L} \) as a map

\[ \mathcal{L} : (\overline{k}[x] \setminus \{0\}) \times (\overline{k}[x] \setminus \{0\}) \to [0, \infty). \]

The local properties plainly carry over to \( \mathcal{L} \):

(i) \( \mathcal{L}(F, G) = \mathcal{L}(G, F) \)

(ii) \( \mathcal{L}(\gamma F, G) = \mathcal{L}(F, G) \) for all \( \gamma \neq 0 \) in \( \overline{k} \)

(iii) \( \mathcal{L}(F_1F_2, G) = \mathcal{L}(F_1, G) + \mathcal{L}(F_2, G) \)

(iv) \( \mathcal{L}(F, G) \geq 0. \)

These last two will prove crucial later on.
Suppose now that \( \tau \) is a \( \mathbb{Q} \)-automorphism of \( \overline{\mathbb{Q}} \) and that \( \alpha, \beta, \tau(\alpha), \tau(\beta) \) are in some field \( L \). Then from lemma 2 and the definition of \( \mathcal{L}_w \) we have

\[
\sum_{\substack{w \in \mathcal{V}_L \\ \text{w} \mid u}} \mathcal{L}_w(x - \alpha, x - \beta) = \sum_{\substack{w \in \mathcal{V}_L \\ \text{w} \mid \mathfrak{p}}} \mathcal{L}_w(x - \tau(\alpha), x - \tau(\beta))
\]

for all places \( v \in \mathcal{V}_Q \), \( v \not\mid \infty \). In particular summing over all the finite places \( v \) we obtain

\[
\mathcal{L}(x - \alpha, x - \beta) = \mathcal{L}(x - \tau(\alpha), x - \tau(\beta)).
\]

From the additivity of \( \mathcal{L} \) we deduce at once the following lemma:

**Lemma 10** For all \( \tau \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \) and all \( F(x) \) and \( G(x) \) in \( \overline{\mathbb{Q}}[x] \setminus \{0\} \)

\[
\mathcal{L}(F, G) = \mathcal{L}(\tau F, \tau G).
\]

We shall need the following key upper bound for \( \mathcal{L} \):

**Lemma 11** Let \( F(x) \) and \( G(x) \) be polynomials in \( k[x] \). Suppose that \( G(x) \) is irreducible in \( k[x] \) and that \( G(x)^e \| F(x) \) (i.e. \( e \) is the highest power of \( G(x) \) which divides \( F(x) \)). Then

\[
\mathcal{L}(F, G) \leq (\partial(F) - e) \log \mu(G) + \partial(G) \log \nu(F) + \partial(G) \log \left( \frac{\partial(F)}{e} \right).
\]

**Proof of lemma 11:** Let \( L \) be the splitting field of \( F(x) \) and \( G(x) \). From the homogeneity of \( \mathcal{L}, \mu \) and \( \nu \) we may as well assume that \( F(x) \) and \( G(x) \) are monic. Further, by the irreducibility of \( G(x) \), each of the linear factors of \( G(x) \) must divide \( F(x) \) to the same power \( e \). In particular, by the additivity of
\( \mathcal{L}, \log \mu \) and \( \partial \), it is enough to consider the case when \( \partial(G) = 1 \). Consequently we set

\[
G(x) = (x - \beta) \\
F(x) = (x - \beta)^e \prod_{i=1}^{L-e} (x - \gamma_i)
\]

where the \( \beta \) and \( \gamma_1, \ldots, \gamma_{L-e} \) are in \( \mathbb{L} \) with \( \gamma_i \neq \beta \). So from the definition of \( \mathcal{L} \)

\[
\mathcal{L}(F, x - \beta) = \sum_{v \in V_L} \sum_{i=1}^{L-e} \left\{ \log^+ |\gamma_i|_v + \log^+ |\beta|_v - \log |\gamma_i - \beta|_v \right\}.
\]

Now

\[
\sum_{i=1}^{L-e} \log |\gamma_i - \beta|_v = \log |\prod_{i=1}^{L-e} (\gamma_i - \beta)|_v = \log |D^e F(\beta)|_v
\]

while by definition of \( \log \mu_v \) and lemma 3

\[
\sum_{i=1}^{L-e} \log^+ |\gamma_i|_v \leq \sum_{i=1}^{L-e} \log^+ |\gamma_i|_v + e \log^+ |\beta|_v \\
= \log \mu_v(F) = \log \nu_v(F)
\]

and

\[
\sum_{i=1}^{L-e} \log^+ |\beta|_v = (\partial(F) - e) \log \mu_v(G).
\]

Hence by the product formula (2.6)

\[
\mathcal{L}(F, x - \beta) \leq \sum_{v \in V_L} \log \nu_v(F) + (\partial(F) - e) \sum_{v \in V_L} \log \mu_v(G) - \sum_{v \in V_L} \log |D^e F(\beta)|_v
\]

\[
= \sum_{v \in V_L} \log \nu_v(F) + (\partial(F) - e) \sum_{v \in V_L} \log \mu_v(G) + \sum_{v \in V_L} \log |D^e F(\beta)|_v.
\]
Applying lemma 5 and recalling that
\[ \sum_{v \in \mathcal{V}_L, v|\infty} \frac{d_v}{d} = 1 \]
we obtain
\[
\sum_{v \in \mathcal{V}_L, v|\infty} \log |D^e F(\beta)|_v \\
\leq \sum_{v \in \mathcal{V}_L, v|\infty} \log \nu_v(F) + \sum_{v \in \mathcal{V}_L, v|\infty} \frac{d_v}{d} \log \left( \frac{\partial(F)}{e} \right) + (\partial(F) - e) \sum_{v \in \mathcal{V}_L, v|\infty} \log^+ |\beta|_v \\
= \sum_{v \in \mathcal{V}_L, v|\infty} \log \nu_v(F) + (\partial(F) - e) \sum_{v \in \mathcal{V}_L, v|\infty} \log \mu_v(G) + \log \left( \frac{\partial(F)}{e} \right). \tag{2.35}
\]
Combining the expressions in (2.34) and (2.35) we conclude that
\[ \mathcal{L}(F, x - \beta) \leq \log \nu(F) + (\partial(F) - e) \log \mu(G) + \log \left( \frac{\partial(F)}{e} \right) \]
as claimed.

It is perhaps worth noting that lemma 11 can also be used to establish results on "automatic vanishing" (see [5]). For example (since \( \log \mu(\Phi_n) = 0 \) for a cyclotomic polynomial \( \Phi_n(x) \)) knowing that
\[ \mathcal{L}(F, \Phi_n) > \phi(n) \log \nu(F) \]
would tell us at once that \( \Phi_n(x) \mid F(x) \).

Finally we determine explicitly the behaviour of \( \mathcal{L} \) on pairs of cyclotomic polynomials:
Lemma 12 Let $\Phi_n(x)$ denote the $n$-th cyclotomic polynomial in $\mathbb{Z}[x]$. Then if $n > m$

$$L(\Phi_n, \Phi_m) = \begin{cases} \phi(m)\Lambda(n/m) & \text{if } m \mid n \\ 0 & \text{otherwise} \end{cases} \quad (2.36)$$

where $\Lambda(n)$ is the Von Mangoldt function. When $n = m$

$$L(\Phi_n, \Phi_n) = \sum_{l \mid n} \{\phi(n) - \phi(n/l)\} \Lambda(l) = \phi(n) \log \left( \frac{n}{\rho(n)} \right) \quad (2.37)$$

where $\rho(n)$ satisfies

$$\log \rho(n) = \sum_{p \mid n} \frac{\log p}{p - 1} \ll \log \log 2n. \quad (2.38)$$

Here, as throughout, $p$ denotes a prime number.

Proof of lemma 12: Let $L$ be a splitting field for $\Phi_n(x)$ and $\Phi_m(x)$. Plainly for any cyclotomic polynomial

$$\log H_v(\Phi_n) = \log \mu_v(\Phi_n) = 0$$

at all the finite places $v$. Hence from lemma 9 and the product formula

$$L(\Phi_n, \Phi_m) = -\sum_{v \in V_L, v \nmid \infty} \log |Res(\Phi_n, \Phi_m)|_v$$

$$= \sum_{v \in V_L, v \mid \infty} \log |Res(\Phi_n, \Phi_m)|_v$$

$$= \log \|Res(\Phi_n, \Phi_m)\|_\infty$$
when $n \neq m$ and

$$\mathcal{L}(\Phi_n, \Phi_m) = - \sum_{v \in \mathcal{V}_L, v \neq \infty} \log |\text{Disc}(\Phi_n)|_v$$

$$= \sum_{v \in \mathcal{V}_L, v \neq \infty} \log |\text{Disc}(\Phi_n)|_v$$

$$= \log ||\text{Disc}(\Phi_n)||_{\infty}$$

when $n = m$. The last equality in each case following from (2.8); since the resultant and discriminant are certainly in $\mathbb{Q}$. The lemma then follows from the familiar formulae for the resultant and discriminant of cyclotomic polynomials:

For $n > m$

$$\text{Res}(\Phi_n, \Phi_m) = \begin{cases} 
    p^{\phi(m)} & \text{if } n = mp^\alpha \text{ with } p \text{ a prime} \\
    1 & \text{else}
\end{cases}$$

and

$$\text{Disc}(\Phi_n) = \begin{cases} 
    (-1)^{\frac{1}{2}\phi(n)} n^{\phi(n)} \prod_{p | n} p^{-\frac{\phi(n)}{p-1}} & \text{for } n > 2 \\
    1 & \text{for } n = 1 \text{ or } 2
\end{cases}$$

Such expressions have been derived by E.T. Lehmer [16] (and more simply by Apostol [1]). Since this lemma is so critical, a proof of these formulae is given below.

From the expression

$$\Phi_n(x) = \frac{x^n - 1}{\prod_{d | n, d < n} \Phi_d(x)} = \frac{x^{n-1} + \cdots + x + 1}{\prod_{d | n, 1 < d < n} \Phi_d(x)}$$

it is not hard to see (for example by induction) that

$$\Phi_n(1) = \begin{cases} 
    0 & \text{if } n = 1 \\
    p & \text{if } n = p^\alpha \text{ } p \text{ a prime and } \alpha \geq 1 \\
    1 & \text{otherwise}
\end{cases} \quad (2.39)$$
Now from the well-known identity
\[ \Phi_m(x) = \prod_{d|m} (x^d - 1)^{\mu(n/d)} \]  
(2.40)
we deduce that when \( n \nmid m \)
\[ \text{Res}(\Phi_n, \Phi_m) = (-1)^{\phi(n)\phi(m)} \prod_{\zeta_n} \Phi_m(\zeta_n) \]
\[ = (-1)^{\phi(n)\phi(m)} \prod_{d|m} \prod_{\zeta_n} (\zeta_n^d - 1)^{\mu(n/d)} \]
\[ = (-1)^M \prod_{d|m \ \text{gcd}(n,d) = s} \left( \prod_{\zeta_n/s} \left(1 - \zeta_n/s\right)^{\phi(n)\mu(m/d)/\phi(d)} \right) \]
(2.41)
\[ = \prod_{d|m \ \text{gcd}(n,d) = s} \Phi_{n/s}(1)^{\phi(n)\mu(m/d)/\phi(d)} \]
since
\[ M = \phi(n) \left( \phi(m) + \sum_{d|m} \mu(m/d) \right) \equiv 0 \pmod{2}. \]
Suppose that \( n = n'D \) and \( m = m'D \) where \( D = \text{gcd}(n, m) \). Now if \( n' \) is not a prime power then \( n/\text{gcd}(n, d) \) is not a prime power for any \( d \mid m \) and by (2.39) the resultant will be 1. So the only cases with \( \text{Res}(\Phi_n, \Phi_m) \neq 1 \) must occur when \( n' = p^\alpha \) for some prime \( p \) and \( \alpha \geq 1 \). In such a case plainly the resultant must be some power of \( p \). Now if \( m \nmid n \) the same argument holds with \( n \) and \( m \) reversed. But clearly \( \text{gcd}(m', p) = 1 \) so \( m' \) cannot be a power of \( p \). Hence the resultant is 1 in this case too. Thus it remains only to consider the case \( n = mp^\alpha \). So by (2.41)
\[ \text{Res}(\Phi_n, \Phi_m) = \prod_{d|m} \Phi_{n/d}(1)^{\phi(n)\mu(m/d)/\phi(d)} \]
\[ = \prod_{d|m} \Phi_{d^\alpha}(1)^{\phi(n)\mu(d)/\phi(d^\alpha)}. \]
Now $\Phi_{dp^\alpha}(1) = 1$ unless $d = p^\gamma$ with $\gamma \geq 0$ while $\mu(p^\gamma) = 0$ unless $\gamma = 0$ or $1$.

Observing the relation

$$\phi(m) = \begin{cases} 
\phi(n)/\phi(p^\alpha) & \text{if } \gcd(m, p) = 1 \\
\phi(n)/\phi(p^\alpha) + \phi(n)\mu(p)/\phi(p^{\alpha+1}) & \text{if } \gcd(m, p) > 1 
\end{cases}$$

then gives us

$$\text{Res}(\Phi_{mp^\alpha}, \Phi_m) = p^{\phi(m)}$$

as claimed.

In the case of the discriminant we see from (2.40) that

$$\Phi_n'(\zeta_n) = n\zeta_n^{n-1} \prod_{d|n, d < n} (\zeta_n^d - 1)^{\mu(n/d)}.$$ 

Hence for $n > 2$

$$\text{Disc}(\Phi_n) = (-1)^{\frac{1}{2}\phi(n)(\phi(n)-1)} \prod_{\zeta_n} \Phi_n'(\zeta_n)$$

$$= (-1)^{\frac{1}{2}\phi(n)n^{\phi(n)}} \prod_{d|n, d < n} \left(\prod_{\zeta_n/d} (1 - \zeta_n^d)^{\phi(n)\mu(n/d)/\phi(n/d)}\right)$$

$$= (-1)^{\frac{1}{2}\phi(n)n^{\phi(n)}} \prod_{d|n, d < n} \Phi_n/d(1)^{\phi(n)\mu(n/d)/\phi(n/d)}$$

since $\phi(n)$ is even. Now by (2.39) the only terms with both $\Phi_n/d(1) \neq 1$ and $\mu(n/d) \neq 0$ occur when $n/d = p$ for some prime $p \mid n$. Hence

$$\text{Disc}(\Phi_n) = (-1)^{\frac{1}{2}\phi(n)n^{\phi(n)}} \prod_{p|n} p^{-\frac{\phi(n)}{p-1}}.$$ 

The results of the lemma then follow since

$$\sum_{l|n} \phi(n/l)\Lambda(l) = \sum_{p^n|n} (\log p) \sum_{i=1}^\alpha \phi(n/p^i)$$
\[= \sum_{p^a \mid n} (\log p) \phi\left(\frac{n}{p^a}\right) \left(1 + (p - 1)(1 + p + \cdots + p^{a-2})\right)
\]
\[= \sum_{p^a \mid n} (\log p) \phi\left(\frac{n}{p^a}\right) p^{a-1}
\]
\[= \phi(n) \sum_{p \mid n} \frac{\log p}{p - 1}
\]

and
\[\sum_{l \mid n} \Lambda(l) = \log n.
\]

The upper bound for \(\rho(n)\) follows simply from
\[\log \rho(n) = \sum_{p \mid n} \frac{\log p}{p - 1}
\]
\[\leq \sum_{p \leq \log n} \frac{\log p}{p} + \frac{\log \log n}{\log n} \sum_{p > \log n} 1 + O(1)
\]
\[= \log \log n + O(1)
\]

using the classical prime number result of Mertens.
Chapter 3

The non-cyclotomic factors

3.1 Proof of theorem 1

As mentioned earlier the proof of theorem 1 relies heavily upon a theorem of Dobrowolski [9]:

**Lemma 13** Given any \( \varepsilon > 0 \) and \( L \geq L_0(\varepsilon) \)

\[
\log \mu(F) \geq \left( \frac{9}{4} - \varepsilon \right) \left( \frac{\log \log L}{\log L} \right)^3
\]

(3.1)

for all irreducible, non-cyclotomic polynomial \( F(x)(\neq x) \) in \( \mathbb{Q}[x] \) with \( \partial(F) \leq L \).

Simpler proofs of this inequality have also been given by Cantor and Straus [6] and by Rausch[20]. The constant \( \frac{9}{4} \) stated here is due to Louboutin [17].

That irreducible \( F(x)(\neq x) \) in \( \mathbb{Q}[x] \) satisfy \( \log \mu(F) = 0 \) if and only if \( F(x) \) is a cyclotomic polynomial was shown by Kronecker. It was conjectured by Lehmer [15] in the 1930’s that in fact any non-zero \( \log \mu(F) \) could be bounded away from zero by a positive constant with

\[
\log \mu(x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1) = 0.162\ldots
\]

46
suggested as an optimal candidate. However Dobrowolski's 1979 bound (itself a great advance over the existing bounds) still remains the best currently available. Dobrowolski, Lawton and Schinzel [11], [10] have also given alternative bounds depending on the number of non-zero coefficients of $F(x)$. Unfortunately these seem inapplicable here. For the subset of non-reciprocal polynomials in $\mathbb{Q}[x]$ we do however have such a bound from a theorem of Smyth [23].

**Lemma 14** Let $F(x)$ be an irreducible, non-reciprocal, non-cyclotomic polynomial in $\mathbb{Q}[x]$. Then

$$\log \mu(F) \geq \log \theta_0 = 0.281\ldots$$

where $\theta_0 = 1.324\ldots$ is the real root of $x^3 - x - 1$.

Recall a polynomial $g(x)$ is non-reciprocal if $x^{\partial(g)}g(\frac{1}{x}) \neq \mp g(x)$.

**Proof of theorem 1.**

We suppose that $F(x)$ factors over $k[x]$ as shown in (1.4). Let $f_i(x)$ denote the minimal polynomial for $\xi_i$ over $k[x]$ and $F_i(x)$ the corresponding minimal polynomial over $\mathbb{Q}[x]$. Then from lemma 6 and the multiplicativity of the Mahler measure $\mu$

$$\mu(F_i) = \mu(x - \xi_i)^{\partial(F_i)}$$

$$\mu(f_i) = \mu(x - \xi_i)^{\partial(f_i)}$$

where

$$\frac{\partial(F_i)}{\partial(f_i)} = \frac{[\mathbb{Q}(\xi_i) : \mathbb{Q}]}{[k(\xi_i) : k]} = \frac{[k : \mathbb{Q}]}{[k(\xi_i) : \mathbb{Q}(\xi_i)]}.$$
In particular
\[
\log \mu(F_i) = \frac{\partial(F_i)}{\partial(f_i)} \log \mu(f_i) = \frac{[k : Q]}{[k(\xi_i) : Q(\xi_i)]} \log \mu(f_i).
\] (3.4)

Now if \( \partial(F_i) \leq r \) with \( r \geq r_0(\varepsilon) \) we see from Dobrowolski's bound that
\[
\log \mu(F_i) \geq \left( \frac{9}{4} - \varepsilon \right) \left( \frac{\log \log r}{\log r} \right)^3
\]
or equivalently
\[
[k(\xi_i) : Q(\xi_i)] \leq \left( \frac{4}{9} + \varepsilon \right) [k : Q] \left( \frac{\log r}{\log \log r} \right)^3 \log \mu(f_i).
\] (3.5)

Using this bound for the factors of small degree gives
\[
\sum_{\partial(F_i) \leq r} m(i) [k(\xi_i) : Q(\xi_i)] \leq \left( \frac{4}{9} + \varepsilon \right) [k : Q] \left( \frac{\log r}{\log \log r} \right)^3 \sum_{\partial(F_i) \leq r} m(i) \log \mu(f_i)
\] (3.6)

where by the non-negativity and additivity of \( \log \mu \) and inequality (2.22)
\[
\sum_{\partial(F_i) \leq r} m(i) \log \mu(f_i) \leq \sum_{i=1}^I m(i) \log \mu(f_i)
\]
\[
= \log \mu(F) \leq \log \nu(F).
\]

While, using (3.3) and the trivial bound, the remaining factors can contribute at most
\[
\sum_{\partial(F_i) > r} m(i) [k(\xi_i) : Q(\xi_i)] \leq \sum_{i=1}^I m(i) [k(\xi_i) : Q(\xi_i)] \frac{\partial(F_i)}{r}
\]
\[
= \frac{[k : Q]}{r} \sum_{i=1}^I m(i) \partial(f_i)
\]
\[
\leq [k : Q] \frac{\partial(F)}{r} \leq [k : Q] \log \nu(F).
\]
Combining these expressions we obtain the desired inequality
\[ \sum_{i=1}^{I} m(i)[k(\xi_i) : Q(\xi_i)] \leq \left( \frac{4}{9} + \varepsilon \right)[k : Q] \left( \frac{\log r}{\log \log r} \right)^3 \log \nu(F) \]
as long as \( r \geq r_0(\varepsilon) \) is suitably large.

Suppose now that \( \xi_i \) is non-reciprocal (in the sense of theorem 1). Then by Smyth's result
\[ \log \mu(F_i) \geq \log \theta_0. \]
Hence by (3.4)
\[ [k(\xi_i) : Q(\xi_i)] \leq [k : Q] \frac{\log \mu(f_i)}{\log \theta_0} \]
for all such \( \xi_i \). It follows then that
\[ \sum_{\xi_i \text{ is non-reciprocal}} m(i)[k(\xi_i) : Q(\xi_i)] \leq \frac{[k : Q]}{\log \theta_0} \sum_{\xi_i \text{ non-reciprocal}} m(i) \log \mu(f_i) \]
\[ \leq [k : Q] \frac{\log \mu(F)}{\log \theta_0} \]
as claimed in (1.19).

### 3.2 Non-cyclotomic bounds and Lehmer's conjecture

In this section we consider the equivalence between best possible upper bounds for the number of non-cyclotomic factors of a polynomial in \( k[x] \) and optimal lower bounds for the Mahler measure of an irreducible, non-cyclotomic polynomial in \( k[x] \).

**Theorem 6** Let \( S \) be a set of irreducible, non-cyclotomic polynomials \( f(x) \) in \( k[x] \) with \( f(0) \neq 0 \). Then the existence of a positive constant \( \theta(S, k) \) satisfying any one of the following statements would imply the truth of the other two:
(I) For all polynomials $f(x)$ in $S$

$$\log \mu(f) \geq \theta(S, k).$$

(3.7)

(II) For all polynomials $F(x)$ in $k[x]$, $F(x)$ factoring in $k[x]$ as shown in (1.4), the number of irreducible factors of $F(x)$ in $S$ satisfies;

$$\sum_{f_i \in S} m(i) \leq \frac{\log \nu(F)}{\theta(S, k)}.$$  

(3.8)

(III) For all polynomials $F(x)$ in $k[x]$, $F(x)$ factoring in $k[x]$ as shown in (1.4), the maximum multiplicity of a factor of $F(x)$ in $S$ satisfies;

$$\max \{m(i) : f_i(x) \in S, 1 \leq i \leq I\} \leq \frac{\log \nu(F)}{\theta(S, k)}.$$  

(3.9)

Clearly then if such a $\theta(S, k) > 0$ exists, each of (I),(II) and (III) must also hold with the plainly optimal constant

$$\beta(S, k) = \inf_{f \in S} \log \mu(f).$$

When

$$S_0 = \{f(x) \in Q[x] : f(x) \text{ is non-reciprocal}\}$$

we have by Smyth's result (3.2) that $\beta(S_0, Q) = \log \theta_0 = 0.281 \ldots$, where $\theta_0 = 1.324\ldots$ is the real zero of $x^3 - x - 1$. Hence inequality (1.19) and the previously claimed optimality of the log $\theta_0$. Now if we let

$$S_k = \{f(x) \in k[x] : f(x) \text{ irreducible, non-cyclotomic, } f(0) \neq 0\}$$

then the Lehmer conjecture amounts to the claim that there exists a positive $\theta(S_k, Q)$ satisfying any of (I),(II) or (III). Now in general for any fields $k \subseteq L$ we have

$$\frac{\beta(S_k, k)}{[L : k]} \leq \beta(S_L, L) \leq \beta(S_k, k).$$
To see this, recall that if $\alpha$ is an algebraic number with minimal polynomial $F(x)$ in $S_L$ and $f(x)$ in $S_k$ then (as in (3.3))

$$\log \mu(F) = \frac{\partial(F)}{\partial(f)} \log \mu(x - \alpha) = \frac{\partial(F)}{\partial(f)} \log \mu(f)$$

where clearly

$$1 \geq \frac{\partial(F)}{\partial(f)} = \frac{[L(\alpha) : L]}{[k(\alpha) : k]} = \frac{[L(\alpha) : k(\alpha)]}{[L : k]} \geq \frac{1}{[L : k]}.$$ 

So in fact Lehmer’s conjecture is equivalent to the existence of a $\theta(S_k, k) > 0$ satisfying (I), (II) or (III) for any $k$. It is believed that when $k = \mathbb{Q}$ we should have $\beta(S_{\mathbb{Q}}, \mathbb{Q}) = \log \theta_2 = 0.162 \ldots$ where $\theta_2 = 1.176 \ldots$ is a real zero of $x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x - 1$. What the correct value should be for a general $k$ is less certain.

It will be immediately apparent from the proof that the theorem might equally well have been stated with the $H(F)$ rather than the $\nu(F)$ height in (II) or (III). Indeed the statement is still quite valid with $\mu(F)$ replacing the $\nu(F)$; though the proof in that case would be essentially a triviality. In fact the only real complexity in the proof of the theorem as stated will occur in proving that (III) implies (I). For this we shall need the following result from Siegel’s lemma:

**Lemma 15** Let $F(x)$ be a polynomial in $k[x]$. Then there exists a polynomial $G(x)$ in $k[x]$ with $\partial(G) < \partial(F)^2$ such that:

$$\log H(FG) = \log \mu(F) + O \left( \log \partial(F) \right) + O(\log c_k) \quad (3.10)$$

$$\log \nu(FG) = \log \mu(F) + O \left( \log \partial(F) \right) + O(\log c_k). \quad (3.11)$$
Here \( c_k = (2/\pi)^{s/d} |\Delta_k|^{1/2d} \) where \( s \) denotes the number of complex places of \( k \), \( d = [k : \mathbb{Q}] \) the degree of \( k \) and \( \Delta_k \) the discriminant of \( k \).

As the proof is non-constructive the nature of \( G(x) \) is not at all clear – although since \( \log \mu(G) = O(\log \partial(F)) \) we can say that such a polynomial ought to contain relatively few non-cyclotomic factors. Now for any \( f(x) \) in \( S \) we can use this lemma to show the existence of a polynomial highly divisible by \( f(x) \) whose height differs little from its Mahler measure. In particular applying the lemma with \( F(x) = (x^3 - x - 1)^m \) and \( k = \mathbb{Q} \) produces the set of extremal polynomials \( Q_m(x) = F(x)G(x) \) in \( \mathbb{Q}[x] \) with

\[
\log H(Q_m) = m \log \theta_0 + O(\log m)
\]

\[
\log \nu(Q_m) = m \log \theta_0 + O(\log m)
\]
described in section 1.3.

**Proof of the lemma:** We shall need to briefly introduce yet another height \( h(F) \) on polynomials in \( \bar{\mathbb{Q}}[x] \). For

\[
F(x) = \sum_{i=0}^{L} a_i x^i \in k[x]
\]

\( h(F) \) is defined by

\[
h(F) = \prod_{v \in \mathbb{V}_k} h_v(F)
\]

where

\[
h_v(F) = \max_{0 \leq i \leq L} |a_i|_v.
\]

Now plainly at the finite places

\[
h_v(F) = H_v(F)
\]
while at the infinite places the two local heights are related by

\[ h_v(F) = \left( \max_{0 \leq i \leq L} |a_i|^2 \right)^{d_v/2d} \]
\[ \leq \left( \sum_{i=0}^{L} |a_i|^2 \right)^{d_v/2d} = H_v(F) \]
\[ \leq (N(F) \max_{0 \leq i \leq L} |a_i|^2)^{d_v/2d} = N(F)^{d_v/2d} h_v(F). \]

Hence summing over all the places we obtain at once the global inequality:

\[ \log h(F) \leq \log H(F) \leq \log h(F) + \frac{1}{2} \log N(F) \]
\[ \leq \log h(F) + \frac{1}{2} \log(\partial(F) + 1). \quad (3.12) \]

Now if \( F(x) \) is a polynomial in \( k[x] \) and \( N \) a positive integer with

\[ M = N - \partial(F) > 0 \]

we can define a set

\[ S_{N,F} = \{ f(x) \in k[x] : \partial(f) < N \text{ and } F(x) | f(x) \}. \]

Then by a result of Bombieri-Vaaler [5, theorem 1] (derived from their form of Siegel’s Lemma [4, theorem 14]) there exist polynomials \( P_1(x), \ldots, P_M(x) \) in \( k[x] \) forming a basis for \( S_{N,F} \) which satisfy

\[ \sum_{i=1}^{M} \log h(P_i) \leq M \log \mu(F) + M \log(c_k) + N^2 u\left(\frac{\partial(F)}{N}\right). \]

The function \( u(\theta) \) has a fairly complicated definition but possesses the simple upper bound for \( 0 < \theta < 1 \)

\[ u(\theta) < \frac{1}{2} \theta^2 \log\left(\frac{1}{4\theta}\right) + \frac{3}{4} \theta^2 \ll \theta^2 \log (2/\theta) \]
which is all that we shall use here. In particular then there is a polynomial $P(x)$ in $S_{N,F}$ with

$$\log h(P) \leq \log \mu(F) + \log c_k + O \left( \frac{(\partial(F))^2}{M} \log \left( \frac{2N}{\partial(F)} \right) \right).$$

(3.13)

So by (3.12) we are guaranteed the existence of a polynomial $G(x) = P(x)/F(x)$ in $k[x]$ with $\partial(G) < M = N - \partial(F)$ and

$$\log H(FG) \leq \log \mu(F) + \log c_k + O \left( \frac{\partial(F)^2}{M} \log \left( \frac{2N}{\partial(F)} \right) \right) + \frac{1}{2} \log N.$$

Choosing $M = \partial(F)^2$ gives us the required upper bound

$$\log H(FG) \leq \log \mu(F) + \log c_k + O (\log \partial(F)).$$

Now from inequality (2.22) and the non-negativity and additivity of $\log \mu$ we always have the lower bound

$$\log H(FG) \geq \log \mu(FG) \geq \log \mu(F)$$

so that (3.10) holds as claimed.

The second expression follows at once on observing that by inequality (2.22)

$$\log H(FG) \leq \log \nu(FG) \leq \log H(FG) + \frac{1}{2} \log N(FG)$$

$$\leq \log H(FG) + \frac{1}{2} \log \left( \partial(F)^2 + \partial(F) \right).$$

**Proof of the theorem:** The proof of $(I) \Rightarrow (II)$ will follow closely that of Schinzel [21, theorem 1]. Suppose that all the polynomials $f(x) \in S$ satisfy (3.7)

$$1 \leq \frac{\log \mu(f)}{\theta(S,k)}$$
and suppose that $F(x) \in k[x]$ factors in $k[x]$ as shown in (1.4). Then by the additivity and non-negativity of $\log \mu$ and the height inequality (2.22) discussed in section 2.2 we have:

$$\sum_{i=1}^{l} m(i) \leq \sum_{i=1}^{l} m(i) \frac{\log \mu(f_i)}{\theta(S, k)} \leq \frac{1}{\theta(S, k)} \sum_{i=1}^{l} \log \mu(f_i^{m(i)}) = \frac{\log \mu(F)}{\theta(S, k)} \leq \frac{\log H(F)}{\theta(S, k)} \leq \frac{\log \nu(F)}{\theta(S, k)}.$$ (II) $\Rightarrow$ (III) is immediate.

Finally we use the above lemma to show that (III) $\Rightarrow$ (I). Let $f(x) \in k[x]$ be a polynomial in $S$. Then for each positive integer $m$ there must be a polynomial $F_m(x) = f(x)^m G_m(x) \in k[x]$ with the comparatively low height

$$\log H(F_m) = \log \mu(f^m) + O(\log(m \vartheta(f))) + O(\log c_k)$$

$$\log \nu(F_m) = \log \mu(f^m) + O(\log(m \vartheta(f))) + O(\log c_k).$$

Since by construction $F_m(x)$ is divisible by $f(x)$ with multiplicity at least $m$ applying condition (III) to $F_m(x)$ gives

$$m \leq \frac{\log \nu(F_m)}{\theta(S, k)} = m \left\{ \frac{\log \mu(f)}{\theta(S, k)} + O\left(\frac{\log(m \vartheta(f))}{m}\right) + O\left(\frac{\log c_k}{m}\right) \right\}.$$}

In other words, for any $\varepsilon > 0$ we can, by taking $m > m_0(\varepsilon, \vartheta(f), c_k)$ suitably large, ensure that

$$1 \leq \frac{\log \mu(f)}{\theta(S, k)} + \varepsilon$$

-condition (I) is therefore plainly assured.
Chapter 4

The cyclotomic factors and theorem 2

4.1 The key inequalities

In this section we prove a number of crucial inequalities for the multiplicities of the cyclotomic factors. We suppose that $F(x)$ factors in $k[x]$ as shown in (1.4). For ease of notation we set

$$a(n) = \frac{1}{\delta(k; n)} \sum_{s=1}^{\delta(k; n)} e(n, s).$$  (4.1)

**Lemma 16** Let $F(x)$ be a polynomial in $k[x]$ whose factorisation over $k[x]$ takes the form (1.4). Then for each positive integer $n$

$$\sum_{m|n} a(m)\phi(m)\Lambda(n/m) + a(n) \sum_{m|n} \{\phi(n) - \phi(n/m)\} \Lambda(m)$$

$$+ \phi(n) \sum_{l=1}^{\infty} a(ln)\Lambda(l)$$

$$\leq \phi(n) \log \nu(F) + \phi(n)a(n) \log \left(\frac{3\delta(F)}{a(n)}\right)$$  (4.2)

where the last term is omitted if $a(n) = 0$.

**Proof of lemma 16:** From lemma 10 it is clear that $L(\Phi_{m,s}, \Phi_n)$ takes the same value for each $s$. In particular by additivity

$$L(\Phi_{m,s}, \Phi_n) = \frac{1}{\delta(k; m)} L(\Phi_m, \Phi_n)$$

56
for $1 \leq s \leq \delta(k; n)$. Hence from the additivity and non-negativity of $\mathcal{L}$ we obtain the lower bound

$$
\mathcal{L}(F, \Phi_n) = \sum_{m=1}^{\infty} \sum_{s=1}^{\delta(k; n)} e(m, s) \mathcal{L}(\Phi_{m,s}, \Phi_n) + \sum_{i=1}^{I} \mathcal{L}(f_i, \Phi_n) \\
\geq \sum_{m=1}^{\infty} a(m) \mathcal{L}(\Phi_m, \Phi_n).
$$

The left-hand side then follows from the explicit expressions for $\mathcal{L}(\Phi_m, \Phi_n)$ given in lemma 12.

The upper bound will come from lemma 11. Since the $\Phi_{n,s}(x)$ are irreducible with $\Phi_{n,s}(x)^{e(n,s)} \| F(x) \) and $\log \mu(\Phi_{n,s}) = 0$ we see that

$$
\mathcal{L}(F, \Phi_n) = \sum_{s=1}^{\delta(k; n)} \mathcal{L}(\Phi_{n,s}, F) \\
\leq \sum_{s=1}^{\delta(k; n)} \left\{ \partial(\Phi_{n,s}) \log \nu(F) + \partial(\Phi_{n,s}) \log \left( \frac{\partial(F)}{e(n, s)} \right) \right\} \\
= \phi(n) \log \nu(F) + \phi(n) \frac{1}{\delta(k; n)} \sum_{s=1}^{\delta(k; n)} \log \left( \frac{\partial(F)}{e(n, s)} \right).
$$

When $e(n, s) \neq 0$ we note the binomial coefficient inequality

$$
\log \left( \frac{\partial(F)}{e(n, s)} \right) \leq e(n, s) \log \left( \frac{3\partial(F)}{e(n, s)} \right)
$$

given for example in [5]. So, by the concavity of the logarithm, when $a(n) \neq 0$ we have

$$
\frac{1}{\delta(k; n)} \sum_{s=1}^{\delta(k; n)} \log \left( \frac{\partial(F)}{e(n, s)} \right) \leq \frac{1}{\delta(k; n)} \sum_{s=1}^{\delta(k; n)} e(n, s) \log \left( \frac{3\partial(F)}{e(n, s)} \right) \\
\leq a(n) \log \left( \frac{3\partial(F)}{a(n)} \right).
$$

The lemma follows on combining these upper and lower bounds for $\mathcal{L}(F, \Phi_n)$. 


Next we derive the inequality we shall need for the theorem 3 bound on the number of $n$-th roots of unity amongst the roots of $F(x)$. For each positive integer $n$ and prime $p$ we define a set

$$T(n,p) = \{u \mid np : u \mid n\}.$$ 

Note that for a fixed $n$ the $T(n,p)$ are plainly disjoint.

**Lemma 17** Let $F(x)$ be a polynomial in $k[x]$ factoring in $k[x]$ as indicated in (1.4). Then for each positive integer $n$ and prime $p$

$$\log p \sum_{m \mid n} a(m)\phi(m) \leq np \log \nu(F) + \sum_{t \in T(n,p)}' a(t)\phi(t) \log \left(\frac{6\delta(F)}{np\alpha(t)}\right)$$

(4.3)

where $\sum'$ indicates a sum in which terms with $a(t) = 0$ are omitted.

**Proof of lemma 17:** We set $b(n)$ to be the left-hand side of (4.2)

$$b(n) = \sum_{m \mid n} a(m)\phi(m)\Lambda(n/m) + a(n)\sum_{m \mid n} \{\phi(n) - \phi(n/m)\}\Lambda(m)$$

$$+ \phi(n)\sum_{l=1}^{\infty} a(ln)\Lambda(l).$$

Then for each positive integer $t$ we have

$$\sum_{u \mid t} b(u) = \sum_{m \mid t} a(m)\phi(m) \sum_{l \mid t/m} \Lambda(l) + \sum_{u \mid t} a(u)\phi(u) \sum_{m \mid u} \Lambda(m)$$

$$- \sum_{m=1}^{\infty} \sum_{l=1}^{\infty} a(ml)\phi(l)\Lambda(m) + \sum_{l=1}^{\infty} \sum_{u=1}^{\infty} a(lu)\phi(u)\Lambda(l)$$

$$= \sum_{u \mid t} a(u)\phi(u) \log(t/u) + \sum_{u \mid t} a(u)\phi(u) \log u$$

$$+ \sum_{l=1}^{\infty} \sum_{u \mid t} a(lu)\phi(u)\Lambda(l)$$

$$= (\log t) \sum_{u \mid t} a(u)\phi(u) + \sum_{l=1}^{\infty} \sum_{u \mid t} a(lu)\phi(u)\Lambda(l).$$
It follows that

\[
\sum_{u \in T(n,p)} b(u) = \sum_{u \mid np} b(u) - \sum_{u \mid n} b(u) \\
= \log(np) \sum_{u \mid np} a(u)\phi(u) + \sum_{l=1}^{\infty} \sum_{u \mid np} a(lu)\phi(u)\Lambda(l) \\
- \log(np) \sum_{u \mid n} a(u)\phi(u) + \log p \sum_{u \mid n} a(u)\phi(u) \\
- \sum_{l=1}^{\infty} \sum_{u \mid n} a(lu)\phi(u)\Lambda(l) \\
= \log(np) \sum_{u \in T(n,p)} a(u)\phi(u) + \log p \sum_{u \mid n} a(u)\phi(u) \\
+ \sum_{l=1}^{\infty} \sum_{u \in T(n,p)} a(lu)\phi(u)\Lambda(l) - \sum_{l=1}^{\infty} \sum_{u \mid np, u \mid n} a(lu)\phi(u)\Lambda(l).
\]

In order to obtain a lower bound for this expression we discard the third sum entirely and note that the fourth sum satisfies

\[
- \sum_{l=1}^{\infty} \sum_{u \mid n} a(lu)\phi(u)\Lambda(l) = - \sum_{m \in T(n,p)} a(m) \sum_{i \mid m} \phi(m/l)\Lambda(l) \\
= - \log p \sum_{m \in T(n,p)} a(m) \sum_{p \mid m, i \geq 1} \phi(m/p^i) \\
= - \log p \sum_{m \in T(n,p)} a(m)\phi(m) \\
\geq - \log 2 \sum_{u \in T(n,p)} a(u)\phi(u).
\]

Therefore

\[
\sum_{u \in T(n,p)} b(u) \geq \log \left( \frac{np}{2} \right) \sum_{u \in T(n,p)} a(u)\phi(u) + \log p \sum_{u \mid n} a(u)\phi(u).
\]
Now using the upper bound in (4.2) and the easily verified identity

$$\sum_{u \in T(n, p)} \phi(u) = \sum_{u \mid n^p} \phi(u) - \sum_{u \mid n} \phi(u) = np - p = n\phi(p)$$

we find that

$$\sum_{t \in T(n, p)} b(u) \leq n\phi(p) \log \nu(F) + \sum_{t \in T(n, p)} a(t)\phi(t) \log \left( \frac{3\partial(F)}{a(t)} \right). \quad (4.5)$$

The lower bound (4.4) and the upper bound (4.5) taken together establish the inequality (4.3) stated in the lemma.

For our bound (1.23) on the number of primitive $n$-th roots of unity we can no longer sum over the divisors of $n$. Instead we shall use the following inequality for a single $a(n)\phi(n)$.

**Lemma 18** Let $F(x)$ be a polynomial in $k[x]$ factoring over $k[x]$ as shown in (1.4). Then for each positive integer $n$ and prime $p$ we have

$$(\log p) a(n)\phi(n) \leq \phi(n)p \log \nu(F) + a(np)\phi(np) \log \left( \frac{3\rho(np)\partial(F)}{npa(np)} \right) \quad (4.6)$$

where the second term on the right is simply omitted if $a(np) = 0$.

**Proof of lemma 18:** From the definition of $\rho(n)$ in lemma 12, inequality (4.2) gives

$$b(np) = \sum_{m \mid np} a(m)\phi(m)\Lambda(n/m) + a(np)\phi(np) \log \left( \frac{np}{\rho(np)} \right)$$

$$+ \phi(np) \sum_{l=1}^{\infty} a(np)\Lambda(l)$$

$$\geq a(n)\phi(n) \log p + a(np)\phi(np) \log \left( \frac{np}{\rho(np)} \right) \quad (4.7)$$
where we have simply discarded the entire third sum and all but one term from
the first. From the upper bound in lemma 16 we see that
\[
\frac{b(np)}{\phi(np)} \leq \frac{\nu(F)}{a(np)} + a(np) \phi(np) \log \left( \frac{3\partial(F)}{a(np)} \right)
\]  
(4.8)

where the second term is omitted whenever \(a(np) = 0\). The lemma follows
immediately from (4.7) and (4.8).

### 4.2 The constant \(c(k)\) and related sums

For the proof of theorem 2 we shall need to know the behaviour of
sums of the form
\[
\sum_{n=1}^{\infty} \delta(k; n)
\]
\[
[k; k] \leq N
\]
as \(N \to \infty\). When \(k = Q\) such sums were considered by Erdős [12] and
Bateman [2].

**Lemma 19 The limits**

\[
c(k) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{\infty} \delta(k; n)
\]
\[
b(k) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{\infty} 1
\]
\[
[k; k] \leq N
\]
exist, with

\[
c(k) = \prod_{p | J} \left( 1 + \frac{1}{p(p-1)} \right) \sum_{m | J} \frac{\delta(k; m)^2 \phi(J/m)}{\phi(m)(J/m)}
\]
\[
b(k) = \prod_{p | J} \left( 1 + \frac{1}{p(p-1)} \right) \sum_{m | J} \frac{\delta(k; m) \phi(J/m)}{\phi(m)(J/m)}
\]
where $J$ is the field constant

$$J = J(k) = \min\{j : k' \subseteq \mathbb{Q}(\zeta_j)\}.$$ 

These quantities possess the following upper bounds

$$c(k) \leq \frac{\zeta(2)\zeta(3)}{\zeta(6)} \min\{[k' : \mathbb{Q}]^2, \ d(J)[k' : \mathbb{Q}], J\}$$

$$b(k) \leq \frac{\zeta(2)\zeta(3)}{\zeta(6)} \min\{d(J), [k' : \mathbb{Q}]\}.$$ 

In the particular case $k = \mathbb{Q}$

$$c(\mathbb{Q}) = b(\mathbb{Q}) = \prod_{p} \left(1 + \frac{1}{p(p-1)}\right) = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \quad (4.9)$$

and for a general cyclotomic field $k = \mathbb{Q}(\zeta_J)$

$$c(\mathbb{Q}(\zeta_J)) = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \prod_{p|J} \left(1 + \frac{1}{p(p-1)}\right)^{-1} \left(1 - \frac{1}{p^2}\right) J. \quad (4.10)$$

**Proof of Lemma 19:** We begin by examining the sums

$$S(x; J, m) = \sum_{\substack{n=1 \atop \phi(n) \leq x \atop \gcd(n, J) = m}} 1.$$ 

Following Bateman [2] we write $S(x; J, m)$ in the form

$$S(x; J, m) = \sum_{n \leq x} a(n; J, m)$$

where

$$a(n; J, m) = \#\{t \in \mathbb{N} : \gcd(t, J) = m \text{ and } \phi(t) = n\}$$
and consider the related Dirichlet's series:

\[
\sum_{n=1}^{\infty} a(n; J, m)n^{-s} = \prod_{p \nmid J} \left( 1 + \sum_{i=1}^{\infty} \phi(p^i)^{-s} \right) \prod_{p^\alpha \parallel J, \beta < \alpha} \phi(p^\beta)^{-s} \prod_{p^\beta \parallel m} \left( \sum_{i=0}^{\infty} \phi(p^{\beta+i})^{-s} \right) \\
= \prod_{p \nmid J} \left( 1 + (p-1)^{-s}(1-p^{-s})^{-1} \right) \phi(m)^{-s} \prod_{p \nmid J/m} \left( 1 - p^{-s} \right)^{-1} \\
= \zeta(s) \prod_{p \nmid J} \left( 1 - p^{-s} + (p-1)^{-s} \phi(m)^{-s} \prod_{p \nmid J/m} \left( 1 - p^{-s} \right) \\
= \zeta(s) F(s; J, m).
\]

Since

\[
| (p-1)^{-s} - p^{-s} | = | s \int_{p-1}^{p} v^{-s-1} dv | \leq | s | (p-1)^{-\Re(s)-1}
\]

it is not hard to see that \( F(s; J, m) \) is an analytic function in the right half-plane. So, by the Weiner-Ikehara theorem [8] :

\[
\lim_{x \to \infty} \frac{S(x; J, m)}{x} = F(1; J, m) = \prod_{p \nmid J} \left( 1 - \frac{1}{p} + \frac{1}{p-1} \right) \frac{1}{\phi(m)} \prod_{p \nmid J/m} \left( 1 - \frac{1}{p} \right) \\
= \prod_{p \nmid J} \left( 1 + \frac{1}{p(p-1)} \right) \frac{1}{\phi(m)} \frac{\phi(J/m)}{(J/m)}. \\
\]

Now from (1.2) and (1.13)

\[
\delta(k; n) = \delta(k; \gcd(n, J))
\]

and

\[
[k(\zeta_n) : k] = \frac{\phi(n)}{\delta(k; \gcd(n, J))}.
\]
Hence we can rewrite the original sum in the form

\[
\sum_{n=1}^{\infty} \delta(k; n) = \sum_{m|J} \delta(k; m) \sum_{\phi(n) \leq N \delta(k; m) \cdot J, m} 1 \sum_{m|J} \delta(k; m) S(N \delta(k; m); J, m).
\]

In particular we see that

\[
c(k) = \lim_{N \to \infty} \sum_{m|J} \delta(k; m) \frac{S(N \delta(k; m); J, m)}{N} = \sum_{m|J} \delta(k; m)^2 \lim_{N \to \infty} \frac{S(N \delta(k; m); J, m)}{N \delta(k; m)} = \sum_{m|J} \delta(k; m)^2 \prod_{p|J} \left(1 + \frac{1}{p(p-1)}\right) \frac{\phi(J/m)}{\phi(m)(J/m)}
\]

as claimed. The expression for \(b(k)\) follows similarly.

The bounds for \(c(k)\) and \(b(k)\) will follow from the simple observation (see (1.3)) that

\[
\delta(k; n) \leq \min\{[k': Q], \phi(n)\}.
\]

Using solely \(\delta(k; n) \leq [k': Q]\) and the multiplicativity of the resulting sum

\[
c(k) \leq \prod_{p|J} \left(1 + \frac{1}{p(p-1)}\right) [k': Q]^2 \sum_{m|J} \frac{\phi(J/m)}{\phi(m)(J/m)} = \frac{\zeta(2)\zeta(3)}{\zeta(6)} [k': Q]^2 \prod_{p|J} \left(1 + \frac{1}{p(p-1)}\right)^{-1} \left(\sum_{i=0}^{\alpha} \frac{\phi(p^\alpha - i)}{\phi(p^i)p^\alpha - i}\right) = \frac{\zeta(2)\zeta(3)}{\zeta(6)} [k': Q]^2 \prod_{p|J} \left(1 + \frac{1}{p(p-1)}\right)^{-1} \left(\frac{p-1}{p} + \sum_{i=1}^{\alpha-1} \frac{1}{p^i} + \frac{1}{p^{\alpha-1}(p-1)}\right) = \frac{\zeta(2)\zeta(3)}{\zeta(6)} [k': Q]^2 \prod_{p|J} \left(1 + \frac{1}{p(p-1)}\right)^{-1} \left(\frac{p-1}{p} + \frac{1}{p-1}\right) = \frac{\zeta(2)\zeta(3)}{\zeta(6)} [k': Q]^2.
\]
Alternatively using $\delta(k; n) \leq [k' : Q]$ for one of the $\delta(k; n)$'s and $\delta(k; n) \leq \phi(n)$ for the other gives

\[ c(k) \leq \prod_{p|J} \left( 1 + \frac{1}{p(p-1)} \right) [k' : Q] \sum_{m|J} \frac{\phi(J/m)}{(J/m)} \]

(4.12)

\[ \leq [k' : Q] \prod_{p} \left( 1 + \frac{1}{p(p-1)} \right) \sum_{m|J} 1 \]

\[ = \frac{\zeta(2)\zeta(3)}{\zeta(6)} d(J)[k' : Q]. \]

(4.13)

Finally using $\delta(k; n) \leq \phi(n)$ in both cases gives

\[ c(k) \leq \prod_{p|J} \left( 1 + \frac{1}{p(p-1)} \right) \sum_{m|J} \frac{\phi(m)\phi(J/m)}{(J/m)} \]

\[ = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \prod_{p|J} \left( 1 + \frac{1}{p(p-1)} \right)^{-1} \prod_{p^a||J} \left( \sum_{i=0}^{\alpha} \frac{\phi(p^i)\phi(p^{\alpha-i})}{p^{\alpha-i}} \right) \]

\[ = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \prod_{p^a||J} \left( 1 + \frac{1}{p(p-1)} \right)^{-1} \left( \frac{p-1}{p} + \sum_{i=1}^{\alpha-1} p^{i-2} - p^{\alpha-1}(p-1) \right) \]

\[ = \frac{\zeta(2)\zeta(3)}{\zeta(6)} J \prod_{p|J} \left( 1 + \frac{1}{p(p-1)} \right)^{-1} \left( 1 - \frac{1}{p^2} \right) \]

(4.14)

\[ \leq \frac{\zeta(2)\zeta(3)}{\zeta(6)} J. \]

(4.15)

The expressions for $b(k)$ follow similarly.

Note that estimates (4.11) and (4.13) can be slightly improved when $k \neq Q$. By the minimality of $J$ the inequality $\delta(k; n) \leq [k' : Q]$ can only have equality when $J$ divides $n$, with $\delta(k; n)$ strictly dividing $[k' : Q]$ otherwise. Similarly we could use multiplicativity to evaluate the sum in (4.12) precisely and gain some improvement on the trivial bound $d(J)$.

Now in the case of a cyclotomic extension $k = Q(\zeta_J)$ we have

\[ \delta(k; n) = [Q(\zeta_J) \cap Q(\zeta_n) : Q] = [Q(\zeta_{\gcd(n,J)}) : Q] = \phi(\gcd(n, J)). \]

(4.16)
In particular we have equality in inequality (4.14) and
\[
c(\Omega(\zeta)) = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \prod_{p \in \mathcal{P}} \left(1 + \frac{1}{p(p - 1)}\right)^{-1} \left(1 - \frac{1}{p^2}\right).
\]

We shall need to apply this lemma in the following form.
\[
\sum_{n=1}^{\infty} \phi(n) = \frac{1}{2}(1 + o(1))c(k)N^2 \tag{4.17}
\]
\[
\sum_{n=1}^{\infty} \left[ k(\zeta_n) : k \right] = \frac{1}{2}(1 + o(1))b(k)N^2. \tag{4.18}
\]

These can be fairly easily obtained from the definitions of \(c(k)\) and \(b(k)\) via partial summation. Let
\[
b(m, k) = \sum_{n=1}^{\infty} \delta(k; n)
\]

then
\[
\sum_{n=1}^{\infty} \phi(n) = \sum_{n=1}^{\infty} \delta(k; n)\left[ k(\zeta_n) : k \right] = \sum_{m \leq N} m b(m, k)
\]

where recall by definition
\[
\sum_{m \leq N} b(m, k) = \sum_{n=1}^{\infty} \delta(k; n) = (1 + o(1))c(k)N.
\]

So applying partial summation (e.g. Hardy and Wright [13, theorem 421])
\[
\sum_{n=1}^{\infty} \phi(n) = N(1 + o(1))c(k)N - \int_{1}^{N} (1 + o(1))c(k)x \, dx
\]
\[
= \frac{1}{2} c(k)N^2(1 + o(1)).
\]

The sum (4.18) follows in the same way.
4.3 A generalisation of theorem 2

In this section we shall prove the following more sophisticated version of theorem 2.

**Theorem 7** Let $F(x)$ be a polynomial in $k[x]$ factoring into irreducibles as indicated in (1.4). Further suppose that $\lambda : \mathbb{N} \to [1, \infty)$ is a function with the following properties

(i) $\lambda(1) = 1$.

(ii) For all $m, n$ in $\mathbb{N}$
\[ \lambda(m) \leq \lambda(mn) \leq n\lambda(m). \]  

(iii) For all $N > N_{\lambda}(\varepsilon)$
\[ M_{\lambda}(N) = \sum_{\lambda(n) \leq N} \phi(n) \leq (1 + \varepsilon) \frac{1}{5\lambda} c_{\lambda} N^2. \]  

Then as long as $\sqrt{\frac{r \log r}{c_{\lambda}}} \geq N_{\lambda}(\varepsilon)$ and $\gg_\varepsilon 1$

\[ \sum_{n=1}^{\infty} a(m) \frac{\phi(m)}{\lambda(m)} \leq (1 + \varepsilon) \sqrt{c_{\lambda}} \partial(F) \left( \frac{\log r}{r} \right)^{\frac{1}{2}}. \]  

The theorem 2 and corollary 1 bounds will arise from the two particular choices of $\lambda$:

\[ \lambda_1(n) = \phi(n) \]
\[ \lambda_2(n) = [k(\zeta_n) : k]. \]
Since \([k(\zeta_m) : k(\zeta_m)] \leq n\) it is clear that both \(\lambda_1\) and \(\lambda_2\) satisfy the first two conditions. From lemma 19 the third condition also holds with

\[
\begin{align*}
    c_{\lambda_1} &= c(Q) = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \\
    c_{\lambda_2} &= c(k)
\end{align*}
\]

for some \(N_{\lambda_1}(\epsilon) = N_0(\epsilon)\) and some \(N_{\lambda_2}(\epsilon) = N_0(\epsilon, k)\). The inequalities then follow, since by definition,

\[
\frac{a(n)\phi(n)}{\lambda_1(n)} = \frac{1}{\delta(k; n)} \sum_{s=1}^{\delta(k; n)} e(n, s)
\]

\[
\frac{a(n)\phi(n)}{\lambda_2(n)} = \sum_{s=1}^{\delta(k; n)} e(n, s).
\]

It is not obvious what other weight functions \(\lambda\) we could usefully take here. Certainly all such \(\lambda\) must satisfy

\[\lambda(n) \leq n\lambda(1) = n.\]

So by standard results (see for example Hardy and Wright [13, theorem 330])

\[M_\lambda(N) \geq \sum_{n \leq N} \phi(n) = \frac{3}{\pi^2} N^2 (1 + o(1))\]

and \(c_\lambda \geq 6/\pi^2\) for any qualifying \(\lambda\).

The weaker condition \(r \geq r_0(\epsilon, \lambda)\) would be enough to prove theorem 2. The form here does however enable more sensitivity to the field. In particular it could be used to replace the condition \(r \geq r_0(\epsilon, k)\) with an explicit \(k\)-dependance. Of course in view of the trivial bound \(\partial(F)\) we could always replace the complicated condition on \(\sqrt{r \log r/c_\lambda}\) by the simpler (if often severer) requirement that \(\log r\) is \(\geq N_\lambda(\epsilon)\) and \(\gg_\epsilon 1\).
Proof of Theorem 7: In addition to the sum $M_\lambda$ of condition (iii) we shall need the following estimate. For all $N \geq N_0(\varepsilon)$ sufficiently large
\[
\log \left( \sum_{n=1 \atop \lambda(n) \leq N}^{\infty} \frac{\phi(m)}{m} \right) \leq \frac{1}{2} \log M_\lambda(N) \left( 1 + \left( \frac{\log \log M_\lambda(N)}{\log M_\lambda(N)} \right) \right)
\]
\[
\leq \frac{1}{2} (1 + \varepsilon) \log M_\lambda(N).
\]
Such a bound can be crudely obtained from the Cauchy-Schwarz Inequality
\[
\sum_{n=1 \atop \lambda(n) \leq N}^{\infty} \frac{\phi(n)}{n} \leq M_\lambda(N)^\frac{1}{2} \left( \sum_{n=1 \atop \lambda(n) \leq N}^{\infty} \frac{\phi(n)}{n^2} \right)^{\frac{1}{2}}
\]
and the rough upper bounds
\[
\sum_{n=1 \atop \lambda(n) \leq N}^{\infty} \frac{\phi(n)}{n^2} \leq \sum_{n \leq M_\lambda(N)} \frac{1}{n} + \sum_{\lambda(n) \leq N} \frac{\phi(n)}{M_\lambda(N)^2} \ll \log M_\lambda(N).
\]

We shall assume throughout that $\varepsilon < 1$ and that
\[
r = \frac{\partial(F)}{\log \nu(F)} \geq 3
\]
since otherwise the result follows trivially. It will now prove convenient to transfer the term
\[
a(n) \sum_{m \mid n} \phi(n) \lambda(m) = a(n) \phi(n) \log(n)
\]
to the right-hand side of (4.2) and to rewrite the resulting inequality in the form
\[
\sum_{m \mid n} a(m) \phi(m) \lambda(n/m) - \sum_{m \mid n} a(n) \phi(n/m) \lambda(m) + \sum_{l=1}^{\infty} a(ln) \phi(n) \lambda(l)
\]
\[
\leq \phi(n) \log \nu(F) + a(n) \phi(n) \log \left( \frac{3\partial(F)}{na(n)} \right). \tag{4.22}
\]
To obtain the required bound we shall sum this inequality over the set of integers

\[ \mathcal{S} = \{ n \in \mathbb{N} : \lambda(n) \leq N \} \]

where \( N \) is a large parameter to be chosen optimally later. For the present we shall simply assume that \( N \geq N_\lambda(\varepsilon) \) and \( N \gg_\varepsilon 1 \) is sufficiently large to ensure that

\[
\sum_{n \in \mathcal{S}} \phi(n) \leq \frac{1}{2} (1 + \frac{\varepsilon}{5}) c_\lambda N^2 \tag{4.23}
\]

\[
\log \left( 3 \sum_{n \in \mathcal{S}} \frac{\phi(n)}{n} \right) \leq (1 + \frac{\varepsilon}{5}) \log(\sqrt{c_\lambda} N). \tag{4.24}
\]

Now it is clear from the second condition that \( nl \in \mathcal{S} \) implies that \( n \in \mathcal{S} \). Such a property enables us to conclude that

\[
\sum_{n \in \mathcal{S}} \left\{ - \sum_{m|n} a(n) \phi(n/m) \Lambda(m) + \sum_{l=1}^{\infty} a(ln) \phi(n) \Lambda(l) \right\} \\
= - \sum_{m=1}^{\infty} \sum_{t=1}^{\infty} a(mt) \phi(t) \Lambda(m) + \sum_{n=1}^{\infty} \sum_{l=1}^{\infty} a(ln) \phi(n) \Lambda(l) \\
= \sum_{n=1}^{\infty} \sum_{l=1}^{\infty} a(ln) \phi(n) \Lambda(l) \geq 0. \tag{4.25}
\]

Meanwhile setting

\[ Y = \sum_{n \in \mathcal{S}} a(n) \phi(n) \]

we still have the trivial bound

\[ Y \leq \sum_{n=1}^{\infty} a(n) \phi(n) = \sum_{n=1}^{\infty} \sum_{s=1}^{\delta(\Phi_{n,s})} e(n, s) \partial(\Phi_{n,s}) \leq \partial(F). \tag{4.26} \]
So by the concavity and monotonicity of the logarithm and bound (4.24) it is readily seen that

\[
\sum_{\substack{n \in S \\ a(n) \neq 0}} a(n) \phi(n) \log \left( \frac{3 \partial(F)}{na(n)} \right) \leq Y \log \left( \sum_{\substack{n \in S \\ a(n) \neq 0}} a(n) \phi(n) \frac{3 \partial(F)}{Y \cdot na(n)} \right) \\
\leq \partial(F) \left( \frac{Y}{\partial(F)} \right) \log \left( \left( \frac{\partial(F)}{Y} \right) 3 \sum_{n \in S} \frac{\phi(N)}{n} \right) \\
\leq \partial(F) \log \left( 3 \sum_{n \in S} \frac{\phi(n)}{n} \right) \\
\leq \left( 1 + \frac{\varepsilon}{5} \right) \partial(F) \log(\sqrt{c_N}N). \tag{4.27}
\]

Similarly by (4.23)

\[
\sum_{n \in S} \phi(n) \log \nu(F) \leq \frac{1}{2} (1 + \frac{\varepsilon}{5}) c_N N^2 \log \nu(F) = \frac{1}{2} (1 + \frac{\varepsilon}{5}) c_N \partial(F) \frac{N^2}{r}. \tag{4.28}
\]

Let \( \Psi(x) = \sum_{n \leq x} \Lambda(n) \). Then by the prime number theorem there exists an \( M = M(\varepsilon) \) such that \( \Psi(x) \geq (1 - \frac{\varepsilon}{5}) x \) whenever \( x \geq M \). Since \( \lambda(lm) \leq l \lambda(m) \), summing the remaining term of inequality (4.22) must therefore contribute at least:

\[
\sum_{n \in S} \sum_{m \mid n} a(m) \phi(m) \Lambda(n/m) = \sum_{m=1}^{\infty} a(m) \phi(m) \sum_{l=1}^{\infty} \Lambda(l) \\
\geq \sum_{m=1}^{\infty} a(m) \phi(m) \sum_{l=1}^{\infty} \Lambda(l) \\
\geq \sum_{\lambda(m) \leq N/M} a(m) \phi(m) \Psi \left( \frac{N}{\lambda(m)} \right) \\
\geq (1 - \frac{\varepsilon}{5}) N \left( \sum_{\lambda(m) \leq N/M} a(m) \phi(m) \frac{\lambda(m)}{\lambda(m)} \right). \tag{4.29}
\]
Hence from (4.25),(4.27),(4.28) and (4.29) summing inequality (4.22) over the integers \( n \) in \( S \) gives

\[
(1 - \frac{\varepsilon}{5})N \sum_{\lambda(m) \leq N/M} \frac{a(m) \phi(m)}{\lambda(m)} \leq (1 + \frac{\varepsilon}{5}) \partial(F) \left\{ \frac{1}{2} c_\lambda \frac{N^2}{r} + \log(\sqrt{c_\lambda} N) \right\}.
\]

That is

\[
\sum_{m=1}^{\infty} \frac{a(m) \phi(m)}{\lambda(m)} \leq (1 + \frac{\varepsilon}{2}) e^\sqrt{c_\lambda} \partial(F) \left\{ \frac{\sqrt{c_\lambda} N}{2r} + \frac{\log(\sqrt{c_\lambda} N)}{\sqrt{c_\lambda} N} \right\}.
\]

Now, assuming that \( N \gg_\varepsilon 1 \) is sufficiently large, we have via the trivial bound (4.26) that

\[
\sum_{m=1}^{\infty} \frac{a(m) \phi(m)}{\lambda(m)} \leq \frac{M}{N} \sum_{m=1}^{\infty} a(m) \phi(m) \leq \partial(F) \frac{M}{N}
\]

\[
\leq \frac{\varepsilon}{2} \partial(F) \frac{\log(\sqrt{c_\lambda} N)}{N}.
\]

So finally, as long as \( N \geq N_\lambda(\varepsilon) \) and \( N \gg_\varepsilon 1 \),

\[
\sum_{m=1}^{\infty} \frac{a(m) \phi(m)}{\lambda(m)} \leq (1 + \varepsilon) e^\sqrt{c_\lambda} \partial(F) \left\{ \frac{\sqrt{c_\lambda} N}{2r} + \frac{\log(\sqrt{c_\lambda} N)}{\sqrt{c_\lambda} N} \right\}.
\]

Choosing

\[
N = \left( \frac{r \log r}{c_\lambda} \right)^{\frac{1}{\varepsilon}}
\]

produces the desired inequality

\[
\sum_{m=1}^{\infty} \frac{a(m) \phi(m)}{\lambda(m)} \leq (1 + \varepsilon) \sqrt{c_\lambda} \partial(F) \left( \frac{\log r}{r} \right)^{\frac{1}{\varepsilon}}
\]

when \( N \) is sufficiently large.
4.4 An alternative proof of the corollary

In section 4.3 we saw that corollary 1 could be obtained from a particular choice of weight function in theorem 7. However it is perhaps better viewed as a statement equivalent to the special case $k = \mathbb{Q}$ of theorem 2. When $k = \mathbb{Q}$ the theorem and corollary are plainly the same. We shall show here how the corollary can be obtained from the particular case $k = \mathbb{Q}$ of the theorem.

Let $\sigma_1, \ldots, \sigma_N$ denote the distinct embeddings of $k$ into $\overline{\mathbb{Q}}$. Then we have

$$
\prod_{i=1}^{N} \sigma_i \Phi_{n,i}(x) = \Phi_n(x)^{N/\delta(k;n)}.
$$

Hence if $F(x)$ is a polynomial in $k[x]$ (factoring in $k[x]$ as shown in (1.4)) we consider the norm

$$
G(x) = \prod_{i=1}^{N} \sigma_i F(x) = g(x) \prod_{n=1}^{\infty} \Phi_n(x)^{N a(n)}
$$

in $\mathbb{Q}[x]$, where $g(x)$ contains the non-cyclotomic factors of $G(x)$ and as before

$$
a(n) = \frac{1}{\delta(k;n)} \sum_{s=1}^{\delta(k;n)} e(n, s).
$$

Plainly

$$
\partial(G) = N \partial(F)
$$

and

$$
\nu(G) \leq \prod_{i=1}^{N} \nu(\sigma_i F) = \nu(F)^N
$$

giving

$$
r_G = \frac{\partial(G)}{\log \nu(G)} \geq \frac{\partial(F)}{\log \nu(F)} = r_F.
$$
So, from the rational case of theorem 2, we see that the number of irreducible factors of \(G(x)\) over \(\mathbb{Q}[x]\) satisfies

\[
\sum_{n=1}^{\infty} Na(n) \leq (1 + \varepsilon) \sqrt{c(\mathbb{Q})} \partial(G) \left( \frac{\log r_G}{r_G} \right)^{\frac{1}{2}}
\leq (1 + \varepsilon) \sqrt{c(\mathbb{Q})} N \partial(F) \left( \frac{\log r_F}{r_F} \right)^{\frac{1}{2}}
\]

as long as \(r_G \geq r_0(\varepsilon, \mathbb{Q})\) is sufficiently large. The result of the corollary follows immediately.

Similarly if \(L\) is any intermediate field \(\mathbb{Q} \subseteq L \subseteq k\) we can use the theorem 2 result for polynomials in \(L[x]\) to give the following weighted inequality:

For any \(\varepsilon > 0\) and \(r > r_0(\varepsilon, L)\) sufficiently large

\[
\sum_{n=1}^{\infty} \frac{1}{[k \cap Q(\zeta_n) : L \cap Q(\zeta_n)]} \sum_{s=1}^{\delta(k,m)} e(n, s) \leq (1 + \varepsilon) \sqrt{c(L)} \partial(F) \left( \frac{\log r}{r} \right)^{\frac{1}{2}}.
\]

This can also be obtained from theorem 7 using the weight function

\[
\lambda_L(n) = [L(\zeta_n) : L].
\]

### 4.5 A generalisation of theorem 3

In this section we shall prove some fairly complicated bounds for the quantities

\[
a(m)\phi(m) = \sum_{s=1}^{\delta(k,m)} e(m, s)\partial(\Phi_{m,s}).
\]

The results of theorem 3 will then follow immediately. Recalling the definition of \(T(n, p)\)

\[
T(n, p) = \{u \mid np : u \mid n\}.
\]
it will be convenient to set
\[ L_n = \frac{1}{\partial(F)} \sum_p \sum_{t \in T(n,p)} a(t) \phi(t) \]
\[ M_n = \frac{1}{\partial(F)} \sum_p a(np) \phi(np) \]
where \( \sum_p \) denotes summation over all primes \( p \). Since for a fixed \( n \) the \( T(n,p) \) are all distinct we see from the trivial bound that
\[ M_n \leq L_n \leq 1. \]

We define two further parameters
\[ R_n = \max \left\{ 4, L_n \left( \frac{r}{n} \right) \log \left( \frac{r}{n} \right) \right\} \]
\[ S_n = \max \left\{ 4, M_n \left( \frac{r}{\phi(n)} \right) \log \left( \frac{r \rho(n)}{\phi(n)} \right) \right\}. \]

**Theorem 8** Let \( F(x) \) be a polynomial in \( k[x] \) factoring in \( k[x] \) as shown in (1.4). Then for each integer \( n \leq r \)
\[ \sum_{m|n} a(m) \phi(m) \ll \partial(F) \left( \frac{n}{r} \right) \left( \frac{R_n}{\log R_n} \right)^{\frac{1}{2}} \quad (4.30) \]
and for \( n \) with \( \phi(n) \leq r \)
\[ a(n) \phi(n) \ll \partial(F) \left( \frac{\phi(n)}{r} \right) \left( \frac{S_n}{\log S_n} \right)^{\frac{1}{2}}. \quad (4.31) \]

The bounds of theorem 3 follow on observing that, since \( L_n \) and \( M_n \) are \( \leq 1 \),
\[ R_n \ll \left( \frac{r}{n} \right) \log \left( \frac{r}{n} \right) \]
\[ S_n \ll \left( \frac{r}{\phi(n)} \right) \log \left( \frac{r \rho(n)}{\phi(n)} \right) \]
where (see (2.38))

$$\log \rho(n) \ll \log \log 20n.$$ 

We shall assume throughout that

$$r = \frac{\partial(F)}{\log \nu(F)} \geq 3$$

since otherwise the inequalities follow automatically from the trivial bound

$$\sum_{m=1}^{\infty} a(m) \phi(m) \leq \partial(F).$$

The first bound (4.30) will come from summing the inequality (4.3)

$$(\log p) \sum_{m|n} a(m) \phi(m) \leq np \log \nu(F) + \sum_{t \in T(n,p)}^{'} a(t) \phi(t) \log \left(\frac{6\partial(F)}{np\alpha(t)}\right)$$

over a set of primes $p \leq N$ where $N \geq 2$ will be chosen optimally later. If $L_n = 0$ then the particular choice $p = 2$ in (4.3) gives

$$\sum_{m|n} a(m) \phi(m) \leq \left(\frac{2}{\log 2}\right) \partial(F) \left(\frac{n}{r}\right)$$

and (4.30) follows immediately. So we may as well assume that $L_n \neq 0$. By the concavity of the logarithm and the definition of $L_n$ we find that

$$\sum_{p \leq N} \sum_{t \in T(n,p)}^{'} a(t) \phi(t) \log \left(\frac{6\partial(F)}{np\alpha(t)}\right) \leq L_n \partial(F) \log \left(\frac{6}{L_n} \sum_{p \leq N} \sum_{t \in T(n,p)}^{'} \frac{\phi(t)}{np}\right)$$

$$\leq L_n \partial(F) \log \left(\frac{6}{L_n} \sum_{p \leq N} 1\right)$$

$$\leq L_n \partial(F) \log \left(\frac{6N}{L_n}\right).$$

From the prime number theorem (and partial summation)

$$\sum_{p \leq N} p \ll \frac{N^2}{\log N} \quad (4.32)$$
while from Chebyshev's Inequality (see for example [13] theorem 414)

\[
\sum_{p \leq N} \log p \geq \frac{\log 2}{4} N
\]

(4.33)

for all \( N \geq 2 \). Hence summing inequality (4.3) over all the primes \( p \leq N \) produces

\[
\sum_{m|n} a(m)\phi(m) \ll \frac{1}{N} \left\{ \left( \frac{N^2}{\log N} \right) n \log \nu(F) + L_n \partial(F) \log \left( \frac{N}{L_n} \right) \right\}
\ll \partial(F) \left\{ \left( \frac{n}{r} \right) \left( \frac{N}{\log N} \right) + \left( \frac{L_n}{N} \right) \log \left( \frac{N}{L_n} \right) \right\}.
\]

Selecting

\[ N = (R_n \log R_n)^{\frac{1}{2}} \]

then produces the bound

\[
\sum_{m|n} a(m)\phi(m) \ll \partial(F) \left( \frac{n}{r} \right) \left( \frac{R_n}{\log R_n} \right)^{\frac{1}{2}} \left\{ 1 + \left( \frac{r}{n} \right) \left( \frac{L_n}{R_n} \right) \log \left( \frac{R_n}{L_n} \right) \right\}.
\]

We consider separately two possible ranges for \( L_n^{-1} \):

Now if

\[ L_n^{-1} \geq \max \left\{ \frac{1}{4} \left( \frac{r}{n} \right) \log \left( \frac{r}{n} \right) , 1 \right\} \]

we have \( R_n = 4 \) and

\[
\left( \frac{r}{n} \right) \left( \frac{L_n}{R_n} \right) \log \left( \frac{R_n}{L_n} \right) \ll \left( \frac{r}{n} \right) \frac{\log(4L_n^{-1})}{L_n^{-2}} \ll 1.
\]

While if

\[ 1 \leq L_n^{-1} < \frac{1}{4} \left( \frac{r}{n} \right) \log \left( \frac{r}{n} \right) \]

then

\[ R_n = L_n \left( \frac{r}{n} \right) \log \left( \frac{r}{n} \right) \]
and
\[
\left( \frac{r}{n} \right) \left( \frac{L_n}{R_n} \right) \log \left( \frac{R_n}{L_n} \right) \ll \frac{\log(R_n/L_n)}{\log(r/n)} \ll 1.
\]

So in either case the required bound
\[
\sum_{m|n} a(m)\phi(m) \ll \partial(F) \left( \frac{n}{r} \right) \left( \frac{R_n}{\log R_n} \right)^{\frac{1}{2}}
\]
holds and (4.30) is proved.

Similarly for (4.31) we average expression (4.6)
\[
(\log p)a(n)\phi(n) \leq \phi(n)p \log \nu(F) + a(np)\phi(np) \log \left( \frac{3\rho(np)\partial(F)}{npa(np)} \right)
\]
over the set of primes \( p \leq N \) for some \( N \geq 2 \). As before, if \( M_n = 0 \) then \( p = 2 \) in the above inequality yields
\[
a(n)\phi(n) \leq \left( \frac{2}{\log 2} \right) \partial(F) \left( \frac{\phi(n)}{r} \right)
\]
and the result is immediate. Hence we shall assume that \( M_n \neq 0 \). This time
\[
\sum_{p \leq N} a(np)\phi(np) \log \left( \frac{3\rho(np)\partial(F)}{npa(np)} \right) \leq \partial(F)M_n \log \left( \frac{3}{M_n} \sum_{p \leq N} \frac{\rho(np)\phi(np)}{np} \right)
\]
\[
\ll \partial(F)M_n \log \left( \frac{N\rho(n)}{M_n} \right). \quad (4.34)
\]
So from (4.32),(4.33) and (4.34) summing the inequality (4.6) over the primes \( p \leq N \) produces a corresponding bound
\[
a(n)\phi(n) \ll \frac{1}{N} \left\{ \left( \frac{N^2}{\log N} \right) \phi(n) \log \nu(F) + \partial(F)M_n \log \left( \frac{N\rho(n)}{M_n} \right) \right\}
\]
\[
\ll \partial(F) \left\{ \left( \frac{\phi(n)}{r} \right) \left( \frac{N}{\log N} \right) + \left( \frac{M_n}{N} \right) \log \left( \frac{N\rho(n)}{M_n} \right) \right\}.
\]

Selecting
\[
N = (S_n \log S_n)^{\frac{1}{2}}
\]
leads to the expression

$$a(n)\phi(n) \ll \partial(F) \left( \frac{\phi(n)}{r} \right) \left( \frac{S_n}{\log S_n} \right)^{\frac{1}{2}} \left\{ 1 + \left( \frac{M_n}{S_n} \right) \left( \frac{r}{\phi(n)} \right) \log \left( \frac{S_n \phi(n)}{M_n} \right) \right\}.$$ 

As in the previous case we consider two possible ranges for $M_n^{-1}$:

If

$$M_n^{-1} \geq \max \left\{ \frac{1}{4} \left( \frac{r}{\phi(n)} \right) \log \left( \frac{r \phi(n)}{\phi(n)} \right), 1 \right\}$$

then we have $S_n = 4$ and

$$\left( \frac{M_n}{S_n} \right) \left( \frac{r}{\phi(n)} \right) \log \left( \frac{S_n \phi(n)}{M_n} \right) \ll \left( \frac{r}{\phi(n)} \right) \frac{\log(4 \phi(n) M_n^{-1})}{M_n^{-1}} \ll 1$$

while if

$$1 \leq M_n^{-1} < \frac{1}{4} \left( \frac{r}{\phi(n)} \right) \log \left( \frac{r \phi(n)}{\phi(n)} \right)$$

then

$$S_n = M_n \left( \frac{r}{\phi(n)} \right) \log \left( \frac{r \phi(n)}{\phi(n)} \right)$$

and

$$\left( \frac{M_n}{S_n} \right) \left( \frac{r}{\phi(n)} \right) \log \left( \frac{S_n \phi(n)}{M_n} \right) \ll \frac{\log(S_n \phi(n)/M_n)}{\log(r \phi(n)/\phi(n))} \ll 1.$$ 

So in both cases we have

$$a(n)\phi(n) \ll \partial(F) \left( \frac{\phi(n)}{r} \right) \left( \frac{S_n}{\log S_n} \right)^{\frac{1}{2}}$$

and the proof of theorem 8 is complete.

### 4.6 Examples for theorem 2

In this section we shall consider examples demonstrating the sharpness of the theorem 2 – at least up to a determination of the precise constant.
All these examples are based upon the family of polynomials

$$F_{N,M}(x) = \prod_{n=1}^{N} \left( x^{nM} - 1 \right)^{N+1-n}. \quad (4.35)$$

When $M = 1$ these polynomials were considered by Dobrowolski [9]. It is readily seen that such polynomials certainly contain a large number of cyclotomic factors. However, as will be shown in the following lemma, they still possess a comparatively small height. In fact we can determine their $\nu$ height exactly.

**Lemma 20** The polynomials $F_{N,M}(x)$ satisfy

$$\partial(F_{N,M}) = MN \left( \frac{N+2}{3} \right) = \frac{1}{6} MN^3 \left( 1 + O\left( \frac{1}{N} \right) \right) \quad (4.36)$$

$$\log \nu(F_{N,M}) = \frac{1}{2} (N+1) \log(N+1) \quad (4.37)$$

$$\log H(F_{N,M}) = \frac{1}{2} N \log N \left( 1 + O\left( \frac{1}{N} \right) \right) \quad (4.38)$$

$$r(F_{N,M}) = \frac{(N+2)NM}{3 \log(N+1)} = \frac{1}{3} MN^2 \left( 1 + O\left( \frac{1}{N} \right) \right). \quad (4.39)$$

**Proof of Lemma 20:** Only properties (4.37) and (4.38) require any justification. In fact, in view of the inequality (2.22)

$$\log H(F_{N,M}) \leq \log \nu(F_{N,M}) \leq \log H(F_{N,M}) + \frac{1}{2} \log N(F_{N,M})$$

where

$$\log N(F_{N,M}) = \log N(F_{N,1}) \leq \log(\partial(F_{N,1}) + 1) \ll \log N$$

it is plainly enough to prove (4.37). Moreover in proving it we may clearly assume that $M = 1$. Now at all the finite places $p$ of $\mathbb{Q}$ we see at once that

$$\log \nu_p(F_{N,1}) = \log \mu_p(F_{N,1}) = 0.$$
Hence
\[ \nu(F_{N,1}) = \nu_\infty(F_{N,1}) = \sup_{\|x\|_\infty = 1} \|F_{N,1}(z)\|_\infty \]

where \(\|\|_\infty\) denotes the usual archimedean absolute value on \(\mathbb{C}\).

Next we let \(V(x)\) denote the \((N+1) \times (N+1)\) matrix \(V(x) = (x^{nm})\) (where \(m = 0, 1, \ldots, N\) indexes rows and \(n = 0, 1, \ldots, N\) columns) and observe that \(F_{N,1}(x)\) can be written in terms of the Vandermonde determinant of \(V(x)\).

\[
\det V(x) = \prod_{l=0}^{N} \prod_{m=0}^{N} \left( x^m - x^l \right) \\
= \prod_{l=0}^{N} \left( \prod_{m=l}^{N} x^l(x^{m-l} - 1) \right) \\
= \prod_{l=0}^{N} \left( \prod_{n=1}^{N} x^l(x^n - 1) \right) \\
= \prod_{n=1}^{N} x^{\frac{1}{2}(N-n+1)(N-n)} (x^n - 1)^{N-n+1} \\
= x^{\frac{1}{2}(N+1)N(N-1)} F_{N,1}(x). 
\]

Now by Hadamard’s Inequality
\[
\nu(F_{N,1}) = \sup_{\|x\|_\infty = 1} \|F_{N,1}(z)\|_\infty = \sup_{\|x\|_\infty = 1} \|\det V(x)\|_\infty \\
\leq (N + 1)^{\frac{1}{2}(N+1)}. 
\]

However for a primitive \((N+1)\)-th root of unity \(\zeta_{N+1}\) the vectors
\[ a_i = (1, \zeta_{N+1}^i, \zeta_{N+1}^{2i}, \ldots, \zeta_{N+1}^{Ni}) \]
satisfy
\[
a_i^* a_j = \sum_{m=0}^{N} \zeta_{N+1}^{m(i-j)} = \begin{cases} 
0 & i \not\equiv j \pmod{N+1} \\
N + 1 & i \equiv j \pmod{N+1}
\end{cases}
\]
where $A^*$ denotes the complex conjugate transpose of $A$. Thus the rows of $V(\zeta_{N+1})$ are orthogonal and we have equality in Hadamard’s Inequality. That is

$$V(\zeta_{N+1})V(\zeta_{N+1})^* = (N + 1)I_{N+1}$$

where $I_{N+1}$ is the $(N + 1) \times (N + 1)$ identity matrix and

$$\| \det V(\zeta_{N+1}) \|_\infty = (N + 1)^{\frac{1}{2}(N+1)}.$$ 

Hence

$$\nu(F_{N,1}) = \sup_{\|z\|_\infty = 1} \| \det V(z) \|_\infty = (N + 1)^{\frac{1}{2}(N+1)}$$

as claimed. This completes the proof of the lemma.

For each possible value of $J(k)$ (i.e. any positive integer $J \neq 2 \pmod{4}$) we shall construct examples where inequality (1.20) is only a small constant factor away from being best possible. We set

$$K = \begin{cases} 2J & \text{if } 2 \nmid J \\ J & \text{if } 4 \mid J \end{cases}$$

and consider the polynomials

$$F_{N,K}(x) = \prod_{n \leq N} (x^{nK} - 1)^{N-n+1}$$

over the field

$$k = \mathbb{Q}(\zeta_J) = \mathbb{Q}(\zeta_K).$$

Clearly $J(k) = J$ for such a field while, as in (4.16),

$$\delta(k; n) = \phi(\gcd(n, J)) = \phi(\gcd(n, K)).$$
Now $F_{N,K}(x)$ factors in $k[x]$ as

$$F_{N,K}(x) = \prod_{n \leq N \atop d | n} \Phi_d(x)^{N-n+1} = \prod_{n \leq N \atop d | n} \prod_{s=1}^{\delta(k,d)} \Phi_{d,s}(x)^{N-n+1}.$$ 

In particular the total number of factors of $F_{N,K}(x)$ satisfies

$$\sum_{n=1}^{\infty} \sum_{s=1}^{\delta(k,n)} e(n, s) = \sum_{n \leq N \atop d | n} \sum_{d \leq N \atop d | n} \delta(k, d)(N - n + 1) = \sum_{d \leq N \atop d | n} \phi(gcd(K, d)) \sum_{n \leq N \atop d | n} (N - n + 1)$$

$$= \sum_{k | K} \phi(k) \sum_{d \leq N \atop d | n} \sum_{n \leq N \atop d | n} \left( N - n + 1 \right)$$

$$= \sum_{k | K} \phi(k) \sum_{m \leq NK/k \atop gcd(m, K/k) = 1} \sum_{m \leq N \atop gcd(m, n) = 1} (N - n + 1).$$

Now the inner sum plainly satisfies

$$\sum_{n \leq N \atop m | n} (N - n + 1) = m \sum_{t \leq N/m} \left( \frac{N + 1}{m} - t \right) = \frac{1}{2} N^2 m + O(N)$$

giving

$$\sum_{n=1}^{\infty} \sum_{s=1}^{\delta(k,n)} e(n, s) = \sum_{k | K} \phi(k) \sum_{m \leq NK/k \atop gcd(m, K/k) = 1} \left( \frac{1}{2} N^2 + O(N) \right)$$

$$= \frac{1}{2} N^2 \sum_{k | K} \phi(k) \left( \sum_{m \leq NK/k \atop gcd(m, K/k) = 1} \frac{1}{m} + O \left( \frac{K}{k} \right) \right).$$

To evaluate this inner sum we note that
\[
\sum_{\substack{m \leq s \\
gcd(m, L) = 1}} \frac{1}{m} = \sum_{m \leq x} \frac{1}{m} \sum_{\frac{m}{s} | m} \mu(s) \\
= \sum_{s | L} \frac{\mu(s)}{s} \sum_{s_n \leq x} \frac{1}{n} s \\
= \sum_{s | L} \frac{\mu(s)}{s} (\log x + O(\log 2s)) \\
= \frac{\phi(L)}{L} \log x + O(d(L)).
\]

Therefore
\[
\sum_{n=1}^{\infty} \sum_{s=1}^{\infty} e(n, s) = \frac{1}{2} N^2 \sum_{k | K} \phi(k) \left\{ \frac{\phi(K/k)}{(K/k)} \log N + O\left( \frac{K}{k} \right) \right\}
\]
\[
= \frac{1}{2} N^2 \log N \left( \sum_{k | K} \frac{\phi(k)\phi(K/k)}{(K/k)} \right) + O(N^2Kd(K))
\]

where by multiplicativity
\[
\sum_{k | K} \frac{\phi(k)\phi(K/k)}{(K/k)} = K \prod_{p | K} \left( 1 - \frac{1}{p^2} \right).
\]

- as was shown in (4.14). So finally the total number of factors of \( F_{N,K}(x) \) over \( k[x] \) must satisfy
\[
\sum_{n=1}^{\infty} \sum_{s=1}^{\infty} e(n, s) = \frac{1}{2} \prod_{p | K} \left( 1 - \frac{1}{p^2} \right) KN^2 \log N \left( 1 + O\left( \frac{d(K)}{\log N} \right) \right). \quad (4.40)
\]

Now by (4.10)
\[
c(k) = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \prod_{p | K} \left( 1 + \frac{1}{p(p-1)} \right)^{-1} \left( 1 - \frac{1}{p^2} \right) K
\]
while by (4.36) and (4.39)
\[
\partial(F_{N,K}) = \frac{1}{6} K N^3 \left( 1 + O\left( \frac{1}{N} \right) \right)
\]
\[
\left( \frac{\log r}{r} \right)^{\frac{1}{2}} = \sqrt{6} \frac{\log N}{K} \left( 1 + O\left( \frac{\log K}{\log N} \right) \right).
\]

Hence as long as \( N \geq N_0(\varepsilon, K) \) is sufficiently large our upper bound (1.20) takes the form
\[
\sum_{n=1}^{\infty} \sum_{s=1}^{\varepsilon(k,n)} e(n, s) \leq \left( 1 + \frac{\varepsilon}{2} \right) \sqrt{c(k)} \partial(F_{N,K}) \left( \frac{\log r}{r} \right)^{\frac{1}{2}} \prod_{p \mid K} \left( \frac{1 - \frac{1}{p^2}}{1 + \frac{1}{p(p-1)}} \right) K N^2 \log N.
\]

Comparison of (4.40) and (4.41) shows the claimed sharpness of theorem 2. Specifically we are off by at most a factor of size
\[
E(J) = \sqrt{\frac{3}{2}} \left( \frac{\zeta(2) \zeta(3)}{\zeta(6)} \right)^{-\frac{1}{2}} \prod_{p \mid K} \left( 1 + \frac{1}{p(p-1)} \right)^{\frac{1}{2}} \left( 1 - \frac{1}{p^2} \right)^{\frac{1}{2}}
\]
\[
= \left( \frac{3 \zeta(6)}{2 \zeta(2) \zeta(3)} \right)^{\frac{1}{2}} \prod_{p \mid K} \left( 1 + \frac{1}{p^2} \right).
\]

In the special case \( k = Q \) we have \( J = 1 \) and \( K = 2 \) and an error factor of size
\[
E(1) = \frac{3}{4} \left( \frac{3 \zeta(6)}{\zeta(2) \zeta(3)} \right)^{\frac{1}{2}} = 0.931 \ldots
\]

Indeed for all \( J \) we have
\[
E(J) \geq E(1) = 0.931 \ldots.
\]

However by taking \( K = \prod_{p \leq x} p \) and letting the number \( x \to \infty \) the accuracy \( E(J) \) increases to
\[
\left( \frac{3 \zeta(6)}{2 \zeta(2) \zeta(3)} \prod_{p} \left( 1 + \frac{1}{p^3} \right) \right)^{\frac{1}{2}} = \left( \frac{3}{2 \zeta(2)} \right)^{\frac{1}{2}} = \left( \frac{3}{\pi} \right) = 0.954 \ldots
\]
Hence as stated earlier we cannot hope in general to improve our bound by any constant factor smaller than $\frac{3}{\pi}$.

4.7 Examples for theorem 8

Next we shall consider examples demonstrating the accuracy of the theorem 8 inequalities. As will be seen this sharpness relies heavily on much more information than is required by our other bounds. In particular significant loss occurs here in going from theorem 8 to theorem 3. Whether there are other extremal examples for which there is also sharpness in theorem 3 remains unclear.

With $F_{N,M}(x)$ as in section 4.6 we shall count the number of $n$-th and primitive $n$-th roots of unity amongst the roots of $F_{N,M}(x)$ for those $n$ with

$$n \mid M.$$  \hspace{1cm} (4.42)

We shall assume that

$$\frac{M}{n} \ll \frac{N}{\log N}.$$  \hspace{1cm} (4.43)

Clearly the particular field $k$ is unimportant here. Hence factoring the $F_{N,M}(x)$ over $\mathbb{Q}[x]$

$$F_{N,M}(x) = \prod_{n \leq N} (x^{nM} - 1)^{N-n+1}$$

$$= \prod_{n \leq N} \prod_{d|nM} \Phi_d(x)^{N-n+1}.$$  

So, since $n \mid M$, we have

$$\sum_{\delta(k,m)} \sum_{m|n} e(m,s) \partial(\Phi_{m,s}) = \sum_{l \leq N} (N - l + 1) \sum_{m|M \atop m|n} \phi(m)$$
\[
= \sum_{i \leq N} (N - l + 1)n \\
= \frac{1}{2} nN(N + 1) \quad (4.44)
\]

and
\[
\sum_{s=1}^{\delta(k,n)} e(n, s) \partial(\Phi_{n,s}) = \sum_{i \leq N \atop i \parallel M} (N - l + 1)\phi(n) \\
= \frac{1}{2} \phi(n)N(N + 1). \quad (4.45)
\]

Now from the definition of \( L_n \)
\[
L_n = \frac{1}{\partial(F_{N,M})} \sum_p \sum_{t \in T(n,p)} a(t)\phi(t) \\
= \frac{1}{\partial(F_{N,M})} \sum_p \sum_{t \in T(n,p)} \phi(t) \sum_{i \leq N \atop i \parallel M} (N - l + 1)
\]

where recall

\[
T(n, p) = \{ u \mid np : u \parallel n \}.
\]

We consider separately those primes \( p \) which do or do not divide \( M/n \). For the primes \( p \mid M/n \) the condition \( t \mid lM \) becomes redundant and
\[
\sum_{p \mid M/n} \sum_{t \in T(n,p)} \phi(t) \sum_{i \leq N \atop i \parallel M} (N - l + 1) = \sum_{p \mid M/n} \sum_{t \in T(n,p)} \phi(t) \sum_{i \leq N} (N - l + 1) \\
= \frac{1}{2} N(N + 1) \sum_{p \mid M/n} \sum_{t \in T(n,p)} \phi(t) \\
= \frac{1}{2} N(N + 1) \sum_{p \mid M/n} n(p - 1) \\
\leq \frac{1}{2} N(N + 1)n \prod_{p \mid M/n} (1 + (p - 1)) \\
\leq \frac{1}{2} N(N + 1)n \left( \frac{M}{n} \right).
\]
For the remaining primes \( p \mid M/n \) the condition \( t \mid lM \) amounts to the restriction \( p \mid l \) and

\[
\sum_{p \mid M/n} \sum_{t \in T(n,p)} \phi(t) \sum_{i \leq N \atop t \mid lM} (N - i + 1) = \sum_{p \leq N} \sum_{t \in T(n,p)} \phi(t) \sum_{i \leq N \atop p \mid i} (N - i + 1) \\
\leq \sum_{p \leq N} \sum_{t \in T(n,p)} \phi(t) \left( \frac{N^2}{p} \right) \\
\leq \sum_{p \leq N} n(p - 1) \left( \frac{N^2}{p} \right) \\
\ll \frac{nN^3}{\log N}.
\]

Hence from condition (4.43)

\[
L_n \ll (MN^3)^{-1} \left\{ N^2M + \frac{nN^3}{\log N} \right\} \\
\ll \frac{n}{M \log N}.
\]

So from the definition of \( R_n \), lemma 20, and (4.43)

\[
R_n = \max \left\{ 4, L_n \left( \frac{r}{n} \right) \log \left( \frac{r}{n} \right) \right\} \\
\ll \left( \frac{n}{M \log N} \right) \left( \frac{MN^2}{n \log N} \right) \log \left( \frac{MN^2}{n \log N} \right) + 1 \\
\ll \frac{N^2}{\log N}
\]

and our bound (4.30) takes the form

\[
\sum_{m \mid n} \sum_{s=1}^{\delta(k,m)} e(m, s) \partial(\Phi_{n,s}) \ll (MN^3) \left( \frac{n}{MN^2 / \log N} \right) \left( \frac{N^2 / \log N}{\log N} \right)^{\frac{1}{2}} \\
\ll nN^2.
\]  

(4.46)

Comparison of (4.44) and (4.46) demonstrates the correct strength of (4.30). Unfortunately in going to the bound (1.22) of theorem 3 we lose a factor of size \( \left( \frac{M}{n \log N} \right)^{\frac{1}{2}} \).
The analysis for (4.31) is very similar, although in addition to (4.42) and (4.43) we shall assume that

\[
\log \left( \frac{\rho(n)M}{\phi(n)} \right) \ll \log N.
\]

From the definition of \( M_n \)

\[
\partial(F_{N,M})M_n = \sum_{p} a(np)\phi(np)
\]

\[
= \sum_{p} \sum_{\substack{l \leq N \\text{np}\|M}} (N - l + 1)\phi(np)
\]

\[
= \sum_{p|M/n} \sum_{l \leq N} (N - l + 1)\phi(np) + \sum_{p|M/n} \sum_{l \leq N} (N - l + 1)\phi(np)
\]

\[
\leq \sum_{p|M/n} N^2\phi(n)p + \sum_{p \leq N} \left( \frac{N^2}{p} \right) \phi(n)p
\]

\[
\ll \phi(n)N^2 \left( \frac{M}{n} \right) + \phi(n)\frac{N^3}{\log N}.
\]

So from condition (4.43) and lemma 20

\[
M_n \ll \frac{\phi(n)}{M \log N}
\]

and

\[
S_n = \max \left\{ 4, M_n \left( \frac{r}{\phi(n)} \right) \log \left( \frac{r \rho(n)}{\phi(n)} \right) \right\}
\]

\[
\ll 1 + \left( \frac{\phi(n)}{M \log N} \right) \left( \frac{MN^2}{\log N} \phi(n) \right) \log \left( \frac{\rho(n) MN^2}{\log N} \phi(n) \right)
\]

\[
\ll \frac{N^2}{\log N}.
\]

Hence finally inequality (4.31) yields

\[
\sum_{s=1}^{s(k,m)} e(n, s) \partial(\Phi_{n,s}) \ll (MN^3) \left( \frac{\phi(n)}{MN^2 / \log N} \right) \left( \frac{N^2 / \log N}{\log N} \right)^{\frac{1}{2}}
\]

\[
\ll \phi(n)N^2.
\]

(4.47)
Comparing (4.45) and (4.47) shows that no significant improvement is possible in (4.31). Again going to inequality (1.23) of theorem 3 unfortunately involves loss of a factor

\[
\left( \frac{M}{\phi(n) \log N} \right)^{\frac{1}{2}}.
\]
Chapter 5

The remaining cyclotomic bounds

5.1 The elementary bounds of theorem 4

When counting only the number of distinct cyclotomic factors the trivial bound $\partial(F)$ is no longer sensible. As observed by Schinzel [21, theorem 3] the most basic upper bounds in this case should be of size essentially $\sqrt{\partial(F)}$. Theorem 4 contains several results of this general form.

Proof of Theorem 4: First, arrange the cyclotomic polynomials $\Phi_{n,s}(x)$ in $k[x]$ in order of increasing degree. Let $P_m(x)$ denote the $m$-th polynomial and $Q_m(x)$ the $m$-th $\Phi_{n,1}(x)$ to appear in the list. Similarly order the positive integers $n$ in terms of increasing $\phi(n)$ with $n_m$ denoting the $m$-th element of this list. Note, when $k = \mathbb{Q}$ we have $P_m(x) = Q_m(x) = \Phi_{n_m}(x)$.

From the definition of $b(k)$ and $c(k)$ it is easy to see that once $\partial(Q_m)$ and $\partial(P_m)$ are $\geq N_0(\varepsilon, k)$

$$m \leq \sum_{n=1}^{\infty} \sum_{[k(\zeta_n):k] \leq \partial(P_m)} \delta(k:n) \leq (1 + \varepsilon)c(k)\partial(P_m).$$

$$m \leq \sum_{n=1}^{\infty} 1 \leq (1 + \varepsilon)b(k)\partial(Q_m).$$

Consequently once $m \geq m_0(\varepsilon, k)$

$$\partial(P_m) \geq (1 - \varepsilon)\left(\frac{m}{c(k)}\right)$$

(5.1)
\[ \partial(Q_m) \geq (1 - \varepsilon) \left( \frac{m}{b(k)} \right). \]  
(5.2)

In particular from the case \( k = Q \)

\[ \phi(n_m) \geq (1 - \varepsilon) \left( \frac{m}{c(Q)} \right) \]
(5.3)

for all \( m \geq m_0(\varepsilon) \) sufficiently large.

Let \( F(x) \) be our polynomial factoring in \( k[x] \) as shown in (1.4) and use \( U, V \) and \( W \) to denote the required sums;

\[
U = \sum_{n=1}^{\infty} \sum_{s=1}^{\infty} \sum_{e(n,s) \neq 0} 1
\]

\[
V = \sum_{n=1}^{\infty} \sum_{e(n,s) \neq 0 \text{ for some } s} 1
\]

\[
W = \sum_{n=1}^{\infty} \frac{1}{\delta(k;n)} \sum_{e(n,s) \neq 0} \delta(k;n)
\]

Then, by (4.17) and (5.1), as long as \( U \geq U_0(\varepsilon, k) \)

\[
\partial(F) \geq \sum_{n=1}^{\infty} \sum_{s=1}^{\infty} \sum_{e(n,s) \neq 0} \partial(\Phi_{n,s}) \geq \sum_{i=1}^{U} \partial(P_i)
\]

\[ \geq \sum_{n=1}^{\infty} \sum_{s=1}^{\infty} \partial(\Phi_{n,s}) \]

\[ \geq \sum_{n=1}^{\infty} \sum_{[k(\zeta_n):k] < \partial(P_U)} \phi(n) \]

\[ \geq \frac{1}{2}(1 - \frac{\varepsilon}{2})c(k)\partial(P_U)^2 \geq \frac{1}{2}(1 - \varepsilon) \left( \frac{U^2}{c(k)} \right). \]

Similarly, from (4.18) and (5.2) when \( V \geq V_0(\varepsilon, k) \)

\[
\partial(F) \geq \sum_{n=1}^{\infty} \sum_{e(n,s) \neq 0 \text{ for some } s} \partial(\Phi_{n,s}) \geq \sum_{i=1}^{V} \partial(Q_i)
\]
\[
\sum_{n=1}^{\infty} \sum_{s=1}^{\delta(k; n)} \partial(\Phi_{n,s}) \geq \sum_{|\mathbf{k}| < \partial(Q_V)} \frac{1}{2}(1 - \frac{\varepsilon}{2}) b(k) \partial(Q_V)^2
\]

\[
\geq \frac{1}{2} (1 - \varepsilon) \left( \frac{V^2}{b(k)} \right).
\]

Hence when \( \partial(F) \geq N_0(\varepsilon, k) \)

\[
U \leq (1 + \varepsilon) \sqrt{2c(k) \partial(F)}
\]

\[
V \leq (1 + \varepsilon) \sqrt{2b(k) \partial(F)}
\]

as claimed.

To prove the weighted version (1.27) let

\[
h(n) = \frac{1}{\delta(k; n)} \sum_{s=1}^{\delta(k; n)} 1
\]

and order the non-zero \( h(n) \) so that

\[
W = \sum_{i=1}^{V} h(m_i)
\]

with

\[
\phi(m_1) \leq \phi(m_2) \leq \cdots \leq \phi(m_V).
\]

So, when \( W > W_0(\varepsilon) \), applying (5.3) and (4.17) gives

\[
\partial(F) \geq \sum_{n=1}^{\infty} \sum_{s=1}^{\delta(k; n)} \partial(\Phi_{n,s})
\]

\[
= \sum_{i=1}^{V} h(m_i) \phi(m_i)
\]

\[
= \sum_{i \leq W} \phi(m_i) + \sum_{i > W} h(m_i) \phi(m_i) - \sum_{i \leq W} (1 - h(m_i)) \phi(m_i)
\]
\[
\sum_{i \leq W} \phi(n_i) + \phi(m_W) \left( \sum_{i \geq W} h(m_i) - \sum_{i \leq W} (1 - h(m_i)) \right) \\
= \sum_{i \leq W} \phi(n_i) + \phi(m_W)(W - [W]) \\
\geq \sum_{n=1}^{\infty} \phi(n) \geq \frac{1}{2} \left( 1 - \frac{\varepsilon}{2} \frac{c(Q)}{\phi(n_W)} \right) \\
\geq \frac{1}{2} \left( 1 - \varepsilon \right) \left( \frac{W^2}{c(Q)} \right).
\]

Hence once \( \partial(F) \geq N_0(\varepsilon) \) we obtain the desired inequality

\[ W \leq (1 + \varepsilon) \sqrt{2c(Q)\partial(F)}. \]

This completes the proof of theorem 4.

The examples

\[ G_{N,k}(x) = \prod_{n=1}^{\infty} \Phi_n(x) \]

\[ [k(\zeta_n):k] \leq N \]

\[ H_{N,k}(x) = \prod_{n=1}^{\infty} \Phi_{n,1}(x) \]

\[ [k(\zeta_n):k] \leq N \]

\[ L_N(x) = \prod_{n=1}^{\infty} \Phi_n(x) \]

\[ \phi(n) \leq N \]

show respectively the sharpness of (1.25),(1.26) and (1.27).

### 5.2 An extension of a theorem of Mann

In this section we shall consider a generalisation of Mann's theorem [18] on linear relations between roots of unity. Given a set of positive integers
our aim here is to determine which primitive $n$-th roots of unity $\zeta_n$ can satisfy a sum of the form

$$\sum_{i=1}^{N} a_i \zeta_n^{n_i} = 0$$

(5.4)

$$0 = n_1 < n_2 < \cdots < n_N$$

with coefficients $a_i$ in some specified algebraic number field $k$ ($a_1 \neq 0$). Following Mann we shall call such a sum 'irreducible' if no subsum vanishes.

**Theorem 9** Suppose $\zeta_n$ satisfies the relation (5.4) above. Then $n$ may be written in the form

$$n = LTW$$

(5.5)

with

(i) $L \mid J$ where $J = J(k)$ is the by now familiar field constant defined in (1.12).

(ii) $T \mid n_i - n_j$ for some pair $i, j \in \{0, 1, \ldots, N\}$ with $i > j$.

(iii) and

$$\frac{W}{d(W)} \leq N - 1.$$ 

In fact $T$ and $W$ can be further subdivided:

$$T = RS$$

$$W = UV$$

where
(iv) $R \mid n_i$ for some $i \in \{2, \ldots, N\}$. Moreover when (5.4) is irreducible $R \mid \gcd(n_2, \ldots, n_N)$.

(v) $n/RL$ is square-free consisting only of primes $p$ with $p \leq N$ and $p \not| J$.

(vi) $TU \mid n_2 \ldots n_N$ and

$$\frac{U}{d(U)} \phi(V) \leq N - 1.$$ 

When $k = Q$ properties (iv) and (v) give us back Mann's original theorem. Basically we are showing that not only must the primes dividing $n$ (not closely linked to exponents $n_i$) be individually less than the number of terms $N$, but in fact by (iii) their product cannot be much larger than this. Of course the cost of this is a much more complicated relationship between the remaining primes and the exponents.

We shall need the following preliminary lemma.

**Lemma 21** For each prime power $p^i$ and coprime integer $m$ we have

$$[k(\zeta_{mp^{i+1}}) : k(\zeta_{mp^i})] = \begin{cases} \frac{p - 1}{p - 1} & \text{if } p \not| J \text{ and } i = 0 \\ \frac{p}{p} & \text{if } p \not| J \text{ and } i \geq 1 \\ \frac{p}{p} & \text{if } p^\alpha \| J \text{ and } i \geq \alpha \geq 1 \end{cases} \quad (5.6)$$

**Proof of the lemma:** Recall from (1.2) that

$$[k(\zeta_n) : k] = \frac{\phi(n)}{\delta(k; n)}$$

where

$$\delta(k; n) = [k \cap Q(\zeta_n) : Q].$$

Hence

$$[k(\zeta_{mp^{i+1}}) : k(\zeta_{mp^i})] = \frac{\phi(mp^{i+1})}{\phi(p^i)} \frac{\delta(k; mp^i)}{\delta(k; mp^{i+1})}.$$
But as remarked earlier (see (1.13))

\[ \delta(k; n) = \delta(k; \gcd(n, J)) \]

and in all the above cases

\[ \gcd(mp^{i+1}, J) = \gcd(mp^i, J). \]

So in each instance

\[ [k(\zeta_{mp^{i+1}}): k(\zeta_{mp^i})] = \frac{\phi(p^{i+1})}{\phi(p^i)} \]

and the result is clear.

**Proof of the theorem:** We proceed by induction on the total number of prime factors \( \Omega(n) \) of \( n \). When \( \Omega(n) = 0 \) we have \( n = 1 \) and the result is vacuously true. Suppose now that \( \Omega(n) > 0 \) and that the theorem is true for all \( m \) with \( \Omega(m) < \Omega(n) \). Without loss of generality we may plainly suppose that (5.4) is irreducible (else we should simply replace it with some subsum that is). In addition we shall assume that

\[ \gcd(n, n_2, \ldots, n_N) = 1. \quad (5.7) \]

For if

\[ R = \gcd(n, n_2, n_3, \ldots, n_N) \neq 1 \]

we have the irreducible sum

\[ \sum_{i=1}^{N} a_i \zeta_{n_i/R} = 0 \]

with

\[ \gcd(n/R, n_2/R, \ldots, n_N/R) = 1 \]
and the result follows by applying the induction hypothesis to $n/R$.

We first show that there must be some $L | J$ such that $n/L$ is square-free and coprime to $J$. That is

(a) If $p^n || J$ then $p^{n+1} | n$

(b) If $p | J$ then $p^2 | n$

in both these cases the proof will be essentially the same. Suppose on the contrary that $n = p^{i+1}m$ where gcd$(m, p) = 1$ and $i \geq 1$ in the case $p | J$ and $i \geq \alpha$ when $p^n || J$. Partitioning the sum into congruence classes $\pmod{p}$ we can rewrite (5.4) as

$$\sum_{r=0}^{p-1} A(r) \zeta_{mp^{i+1}}^r = 0$$

where

$$A(r) = \sum_{\substack{j=1 \quad N \quad a_j \zeta_{mp^{i+1}}^{n_j-r} \quad n_j \equiv r \pmod{p}}}$$

Now plainly each of the $A(r)$ are in $k(\zeta_{mp^i})$ producing a polynomial

$$P(x) = \sum_{r=0}^{p-1} A(r)x^r$$

in $k(\zeta_{mp^i})[x]$, of degree at most $p - 1$ vanishing at $\zeta_{mp^{i+1}}$. But from the lemma we have

$$[k(\zeta_{mp^{i+1}}) : k(\zeta_{mp^i})] = p.$$ 

So $P(x)$ must be the identically zero polynomial, all the $A(r)$ vanish and

$$\sum_{\substack{j=1 \quad N \quad a_j \zeta_{mp^{i+1}}^{n_j} \quad n_j \equiv r \pmod{p}}}$$
for each $0 \leq r \leq p - 1$. From the irreducibility of our original sum the $n_i$ must therefore all lie in the same residue class (mod $p$). In particular, since $n_1 = 0$, all the $n_i \equiv 0 \pmod{p}$. But this contradicts our assumption that $\gcd(n, n_2, \ldots, n_N) = 1$. Hence no such configurations can occur; as stated in (v).

We suppose next that $p \mid n/L$. From our previous comments we know that $p^2 \mid n$ and $p \mid J$. Therefore writing $n = mp$ with $\gcd(m, p) = 1$ we may set

$$\zeta_n = \zeta_m \zeta_p$$

where $\zeta_m$ and $\zeta_p$ are some primitive $m$-th and $p$-th roots of unity. As before, dividing the sum (5.4) into residue classes (mod $p$) we obtain

$$\sum_{r=0}^{p-1} B(r) \zeta_p^r = 0$$

where

$$B(r) = \sum_{\substack{i=1 \n_i \equiv r \pmod{p}}}^N a_i \zeta_m^{n_i}.$$  

Clearly each $B(r)$ is in $k(\zeta_m)$ giving us a polynomial

$$Q(x) = \sum_{r=0}^{p-1} B(r) x^r$$

in $k(\zeta_m)[x]$ of degree at most $p - 1$ vanishing at $\zeta_p$. But from the lemma

$$[k(\zeta_m)(\zeta_p) : k(\zeta_m)] = p - 1.$$  

In particular the minimal polynomial for $\zeta_p$ over $k(\zeta_m)$ must be just the familiar polynomial

$$\Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + x + 1.$$
Now \( Q(x) \) cannot be an identically zero polynomial since (exactly as in the previous analysis) this would contradict either the irreducibility of (5.4) or the coprimeness (5.7) of \( n \) and the exponents \( n_i \). Hence \( Q(x) \) is some non-zero multiple of \( \Phi_p(x) \) and

\[
B(0) = B(1) = \cdots = B(p-1) \neq 0.
\]

As an immediate consequence we see that the number of terms \( N \) must have been at least as great as the number of residue classes \( p \). That is \( p \leq N \) as claimed in (v).

As is suggested by condition (v) the primes \( p \nmid n_2 \cdots n_N \) will be slightly more straightforward to deal with. Suppose that \( p \mid n/L \) with \( p \nmid n_2 \cdots n_N \). Then, since none of the \( n_i \equiv 0 \pmod{p} \) for \( 2 \leq i \leq N \), the sum \( B(0) \) consists of the single term

\[
B(0) = a_1.
\]

So the remaining \( N-1 \) exponents are partitioned amongst the remaining \( p-1 \) residue classes. In particular we are assured of a residue class \( r \neq 0 \) with

\[
\sum_{i=2}^{N} \frac{1}{n_i \equiv r \pmod{p}} \leq \frac{N-1}{p-1}.
\]

Thus, from the relation \( B(0) = B(r) \), we obtain a new shorter sum

\[
a_0 - \sum_{\substack{i=1 \leq N \\text{and} \ n_i \equiv r \pmod{p}}} a_i \zeta_m^{n_i} = 0
\]

with at most

\[
\left( \frac{N-1}{p-1} \right) + 1
\]
terms. By the induction hypothesis this \( m \) certainly satisfies

\[
m = LT'(U'V')
\]

with

\[
T' \mid n_i - n_j
\]

for some \( 0 \leq j < i \leq N \)

\[
T'U' \mid n_2 \ldots n_N
\]

and

\[
\frac{U'}{d(U')} \phi(V') \leq \frac{N-1}{p-1}.
\]

Now since \( \gcd(m, p) = 1 \) we may write this latter condition in the form

\[
\frac{U'}{d(U')} \phi(V'p) \leq N - 1.
\]

Hence we can take

\[
n = L(RS)(UV)
\]

with

\[
R = 1, \ S = T', \ U = U', \ V = V'p.
\]

From our previous discussions concerning (v) such a representation is easily seen to satisfy requirements (i) through (vi).

There remains now only to consider the case when all the primes \( p \mid n/L \) divide \( n_2 \ldots n_N \). In this case \( B(0) \) will no longer consist of a single term, so instead we shall pick the two residue classes \( r_1 \) and \( r_2 \) containing the fewest exponents \( n_i \pmod{p} \).
Let

\[ C(r) = \sum_{n_i \equiv r \pmod{p}} 1 \]

then summing over all possible pairs of residue classes

\[ N(p - 1) = \sum_{r=0}^{p-1} \sum_{s=0}^{p-1} \{C(r) + C(s)\} \geq \frac{1}{2} p(p - 1) \{C(r_1) + C(r_2)\}. \]

Hence plainly

\[ C(r_1) + C(r_2) \leq \frac{2N}{p} \]

and relation \( B(r_1) = B(r_2) \) leads to a sum

\[ \sum_{n_i \equiv r_1 \pmod{p}}^{N} a_i \zeta_m^{n_i} - \sum_{n_j \equiv r_2 \pmod{p}}^{N} a_j \zeta_m^{n_j} \]

with at most \( 2N/p \) terms. Setting

\[ n_I = \min \{n_i : n_i \equiv r_1 \text{ or } r_2 \pmod{p}\} \]

leads finally to a sum of the form (5.4)

\[ a_I + \sum_{n_i \equiv r_1 \pmod{p}}^{N} a_i \zeta_m^{n_i - n_I} - \sum_{n_j \equiv r_2 \pmod{p}}^{N} a_j \zeta_m^{n_j - n_I} = 0 \]

containing no more than \( 2N/p \) elements. So by the induction hypothesis we may write

\[ m = L T' W' \]

with

\[ T' | (n_j - n_I) - (n_i - n_I) = n_j - n_i \]

for some \( 0 \leq i < j \leq N \), and

\[ \frac{W'}{d(W')} \leq \frac{2N}{p} - 1. \]
Since \( \gcd(m, p) = 1 \) we see from this last condition that

\[
\frac{W'p}{d(W'p)} \leq N - \frac{p}{2} \leq N - 1.
\]

Hence we may write

\[ n = L(RS)(UV) \]

with

\[ R = 1, \ S = T', \ U = W'p, \ V = 1. \]

Since we have already shown that \( n/L \) satisfies (v) and since by assumption all the primes dividing \( n/L \) divide \( n_2 \ldots n_N \) it is not too difficult to see that this representation does indeed have the properties (i) through (v).

This completes the proof of the theorem.

### 5.3 A generalisation of theorem 5

One particularly noticeable aspect of theorem 9 is the irrelevancy of the coefficients \( a_i \). It is no suprise then that our theorem 5 bound is really a bound on the number of cyclotomic factors of a whole set of polynomials (who happen to share the same exponents) rather than of an individual polynomial.

Consequently, given a vector \( \vec{n} = (n_1, n_2, \ldots, n_N) \) in \((\mathbb{N} \cup \{0\})^N\) we define

\[
E = E(\vec{n}) = n_2 \cdots n_N
\]

\[
G(\vec{n}) = \left\{ d \mid E : d = RS \text{ where } R \mid n_i \text{ some } i \geq 2 \text{ and } S \prod_{p \leq N} p \right\}
\]

\[
\Delta(\vec{n}) = \# \{ d \in G(\vec{n}) : d \mid n_i - n_j \text{ some } 1 \leq j < i \leq N \} \prod_{\substack{p \leq N \mid E}} \left(1 + \frac{1}{p}\right)
\]
and the set of polynomials

\[ \mathcal{F}(\vec{n}) = \left\{ f(x) = \sum_{i=1}^{N} a_i x^{n_i} : a_i \in k, \text{ with } a_1 \neq 0 \right\}. \]

We then have the following theorem

**Theorem 10** Let \( \vec{n} = (n_1, \ldots, n_N) \) be a vector in \((\mathbb{N} \cup \{0\})^N\) with \( n_1 = 0 \). Then

\[ \# \{ n \in \mathbb{N} : f(\zeta_n) = 0 \text{ for some } f(x) \in \mathcal{F}(\vec{n}) \text{ and some } \zeta_n \} \ll d(J) \Delta(\vec{n}) N. \]

Here \( \zeta_n \) denotes a primitive \( n \)-th root of unity and \( J = J(k) \) is the familiar field constant \((1.12)\).

The theorem 5 bound for a single polynomial in \( \mathcal{F}(\vec{n}) \) follows immediately.

The proof of theorem 10 will simply involve counting all the possible \( n \) which satisfy the conditions (i) through (vi) of theorem 9. We shall need a small arithmetic lemma:

**Lemma 22** Let \( g : \mathbb{N} \to [1, \infty) \) be a non-negative multiplicative function with

\[ \sum_{g(p) \leq Y} \log p \ll Y \]

for all \( Y \geq 1 \). Then the number of square-free integers \( m \) with \( g(m) \leq M \) satisfies

\[ \sum_{m=1}^{\infty} |\mu(m)| \ll \frac{M}{\log M} \prod_{g(p) \leq M} \left( 1 + \frac{1}{g(p)} \right) \]

for all \( M \geq 2 \).
**Proof of lemma 22:** Since we are only aiming for a rough upper bound (rather than an asymptotic result) the proof is fairly simple.

\[
\sum_{m=1}^{\infty} |\mu(m)| \leq \sum_{m=\sqrt{M}}^{\infty} \frac{|\mu(m)|}{\log M} \frac{2 \log m}{\log M} + \sqrt{M}
\]

\[
\leq \frac{2}{\log M} \sum_{m=1}^{\infty} |\mu(m)| \sum_{p|m} \log p + \sqrt{M}
\]

\[
= \frac{2}{\log M} \sum_{m=1}^{\infty} \sum_{p} |\mu(mp)| \log p + \sqrt{M}
\]

\[
\leq \frac{2}{\log M} \sum_{m=1}^{\infty} |\mu(m)| \sum_{p, g(p) \leq M/\sqrt{M}} \log p + \sqrt{M}
\]

\[
\ll \frac{M}{\log M} \sum_{m=1}^{\infty} \frac{|\mu(m)|}{g(m)}
\]

\[
\leq \frac{M}{\log M} \prod_{g(p) \leq M} \left(1 + \frac{1}{g(p)}\right)
\]

as claimed.

**Proof of Theorem 10:** From Theorem 9

\[
\mathcal{N}(\bar{n}) = \# \{n \in \mathbb{N} : f(\zeta_n) = 0 \text{ for some } f(x) \in \mathcal{F} \bar{n} \text{ and some } \zeta_n\}
\]

satisfies

\[
\mathcal{N}(\bar{n}) \leq \sum_{n=1}^{\infty} 1
\]

where \(L, R, S, U\) and \(V\) satisfy the conditions of theorem 9. Hence

\[
\mathcal{N}(\bar{n}) \leq \left(\sum_{L|J} 1\right) \left(\sum_{d \in \varphi(n)} 1\right) \left(\sum_{m \in \mathcal{M}} |\mu(m)|\right)
\]
where
\[ \mathcal{M} = \left\{ m = UV : U \mid n_2 \cdots n_N \text{ and } \frac{U}{d(U)} \phi(V) \leq N \right\}. \]

Now if \( g(m) \) is the totally multiplicative function generated by
\[
g(p) = \begin{cases} 
p - 1 & \text{if } p \nmid n_2 \cdots n_N \\
\frac{1}{2}p & \text{if } p \mid n_2 \cdots n_N \end{cases}
\]
then \( g(m) \) satisfies the conditions of lemma 22 and
\[
\sum_{m \in \mathcal{M}} |\mu(m)| \leq \sum_{m=1}^{\infty} |\mu(m)| \leq \frac{N}{\log N} \prod_{p \leq N+1} \left( 1 + \frac{1}{p-1} \right) \prod_{p \leq 2N} \left( 1 + \frac{2}{p} \right)
\]
\[
\leq \frac{N}{\log N} \prod_{p \leq N} \left( 1 + \frac{1}{p} \right) \prod_{p \leq N} \left( 1 + \frac{1}{p} \right)
\]
\[
\leq N \prod_{p \leq N} \left( 1 + \frac{1}{p} \right).
\]

Hence
\[
\mathcal{N}(\bar{n}) \leq d(J) \left( \sum_{d \in \mathcal{O} \setminus \{0\}} 1 \right) N \prod_{p \leq N} \left( 1 + \frac{1}{p} \right)
\]
\[
= d(J) \Delta(\bar{n}) N
\]
as claimed.

To complete the theorems there remains only to justify the claim (1.30) that \( F(x) \) cannot have a factor of multiplicity greater than its number of non-zero coefficients \( N(F) \).
Lemma 23 Suppose
\[ F(z) = \sum_{j=1}^{N(F)} a_j x^{n_j} \in \mathbb{C}[x] \]
has a zero \( \alpha \neq 0 \) of multiplicity greater than or equal to \( N(F) \). Then \( F(x) \equiv 0 \).

Such a lemma can be found in Schinzel-Montgomery [19, lemma 2] and no doubt elsewhere too.

Proof of lemma 23: First we may assume (after multiplying through by some power of \( x \) if necessary) that all the \( n_j \geq N(F) \). We write
\[ \mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_{N(F)} \end{pmatrix} \]
for the vector of coefficients and \( C \) for the matrix
\[ C = \begin{pmatrix} (n_j^i) \\ \vdots \\ (a_{N(F)}^i) \end{pmatrix} \]
where \( 0 \leq i \leq N(F) - 1 \) indexes rows and \( 1 \leq j \leq N(F) \) indexes columns.

From the relations
\[ \alpha^i F^i(\alpha) = 0 \quad 0 \leq i \leq N(F) - 1 \]
we see that
\[ C \mathbf{a} = 0. \]

But writing the determinant of \( C \) in terms of a Vandermonde determinant
\[ |\det C| = \left( \prod_{j=1}^{N(F)} |\alpha^{n_j}| \right) |\det \begin{pmatrix} (n_j^i) \\ \vdots \\ (a_{N(F)}^i) \end{pmatrix}| \]
\[ = \left( \prod_{j=1}^{N(F)} |\alpha^{n_j}| \right) |\det \begin{pmatrix} n_j^i \end{pmatrix}| \]
\[
\left( \prod_{j=1}^{N(F)} |\alpha^{n_j}| \right)^{N(F)} \prod_{j=1}^{N(F)} \prod_{k=1 \atop j<k}^{n_k - n_j} \neq 0.
\]

So $C$ is non-singular

\[
a = C^{-1} \mathbf{0} = \mathbf{0}
\]

and

\[
F(x) \equiv 0
\]

as claimed.

This completes the proof of the theorems.
Appendix A

Intersections of cyclotomic fields

Let $\zeta_n$ and $\zeta_m$ be primitive $n$-th and $m$-th roots of unity. Then

$$Q(\zeta_n) \cap Q(\zeta_m) = Q(\zeta_{\gcd(n,m)})$$

Since the proof is very simple we give it here:

**Proof:** It is not hard to see that the compositum

$$Q(\zeta_n)Q(\zeta_m) = Q(\zeta_{[n,m]})$$

where $[n,m]$ denotes the least common multiple of $n$ and $m$. Hence we have the galois extensions

```
    Q(\zeta_{[n,m]})
      /   \
     /     \
Q(\zeta_n) - Q(\zeta_m)
      /   \
     /     \
Q(\zeta_n) \cap Q(\zeta_m)
          /   \
         /     \
          Q   
```
with

\[
\frac{[Q(\zeta_n) \cap Q(\zeta_m) : Q]}{[Q(\zeta_m) : Q(\zeta_n) \cap Q(\zeta_m)]} = \frac{[Q(\zeta_m) : Q]}{[Q(\zeta_n, m) : Q(\zeta_n)]}
\]

\[
= \frac{[Q(\zeta_n) : Q][Q(\zeta_m) : Q]}{[Q(\zeta_{n,m}) : Q]}
\]

\[
= \frac{\phi(n)\phi(m)}{\phi([n, m])}
\]

\[
= \phi(\gcd(n, m)).
\]

But clearly

\[
Q \subseteq Q(\zeta_{\gcd(n,m)}) \subseteq Q(\zeta_n) \cap Q(\zeta_m)
\]

so by comparison of degrees

\[
Q(\zeta_{\gcd(n,m)}) = Q(\zeta_n) \cap Q(\zeta_m)
\]

as claimed.
Appendix B

Properties of $J$

Let $k$ be an algebraic number field $k'$ its maximal abelian subfield and

$$J = J(k) = \min \{ j : k' \subseteq \mathbb{Q}(\zeta_j) \}.$$  

Suppose that $J$ factors as

$$J = \prod_{i=1}^{m} p_i^{\alpha_i}$$

and set $M$ to be

$$M = 2^r \prod_{i=1}^{m} p_i \quad \varepsilon = \begin{pmatrix} 0 & 2J \\ 1 & 2J \end{pmatrix}.$$  

Then

$$[k' : \mathbb{Q}] = \frac{J}{M} [k' \cap \mathbb{Q}(\zeta_M) : \mathbb{Q}]. \quad (B.1)$$

In particular since the primes dividing $J$ are exactly those that ramify and hence those that divide the discriminant $\Delta_{k'}$ of $k'$ (see a proof of Kronecker-Weber e.g [25]) we obtain the following bound on $d(J)$:

$$d(J) \leq \frac{3}{2} d([k' : \mathbb{Q}]) 2^{w(\Delta_{k'})}$$

where $w(n)$ denotes the number of distinct prime factors of $n$.

We shall need the following simple lemma:
Lemma 24. Let $L$ be a number field. If $p$ is a prime with

$$L(\zeta_{p^\alpha}) = L(\zeta_{p^{\alpha-1}}) \quad \alpha \geq 2 \quad p \neq 2$$

$$\quad \alpha \geq 3 \quad p = 2$$

then

$$L(\zeta_{p^\alpha}) = \begin{cases} 
L(\zeta_p) & p \neq 2 \\
L(i) & p = 2.
\end{cases}$$

Proof of the lemma: Plainly it is enough to show that

$$[L(\zeta_{p^\beta}) : L(\zeta_{p^{\beta-1}})] = p \quad \beta \geq 3 \quad p \neq 2$$

$$\quad \beta \geq 4 \quad p = 2$$

implies that

$$[L(\zeta_{p^\beta}) : L(\zeta_{p^{\beta-1}})] = p$$

$$[L(\zeta_{p^{\beta-1}}) : L(\zeta_{p^{\beta-2}})] = 1.$$ 

We consider first the case when $p \neq 2$. Since $[L(\zeta_{p^\beta}) : L(\zeta_{p^{\beta-2}})] = p$

we see that

$$Gal(L(\zeta_{p^\beta})/L(\zeta_{p^{\beta-2}})) = \{\sigma, \sigma^2, \ldots, \sigma^p = Identity\}$$

where

$$\sigma : \zeta_{p^\beta} \mapsto \zeta_{p^\beta}^a$$

for some positive integer $a$. Since $\sigma^a = Identity$ this $a$ must satisfy

$$a^p \equiv 1 \pmod{p^\beta}.$$ 

But when $p \neq 2$ this implies that:

$$a \equiv 1 \pmod{p^{\beta-1}}.$$
To see this write $a = 1 + up$ (by "little Fermat" $a \equiv a^p \equiv 1 \pmod{p}$) and observe that (since $p \neq 2$)
\[
a^{p-1} + a^{p-2} + \cdots + a + 1 \equiv \sum_{i=0}^{p-1} (1 + iup) \pmod{p^2}
= p(1 + \frac{1}{2}up(p - 1)) \equiv p \pmod{p^2}.
\]
So $p^2 \mid (a^{p-1} + \cdots + a + 1)$ and as claimed
\[
p^\beta | a^p - 1 = (a - 1)(a^{p-1} + \cdots + a + 1) \Rightarrow p^{\beta-1} | (a - 1).
\]
Now $a \equiv 1 \pmod{p^{\beta-1}}$ implies that $\zeta_{p^{\beta-1}}$ is fixed by $Gal(L(\zeta_{p^{\beta}})/L(\zeta_{p^{\beta-1}}))$.
Since this extension is clearly galois we see that $\zeta_{p^{\beta-1}}$ is in $L(\zeta_{p^{\beta-2}})$ and the result is immediate.

The case $p = 2$ is handled more directly. Since
\[
x^4 - \zeta_{2^{\beta-2}} = (x + \zeta_{2^\beta})(x - \zeta_{2^\beta})(x + i\zeta_{2^\beta})(x - i\zeta_{2^\beta})
\]
the minimum polynomial for $\zeta_{2^\beta}$ must be one of
\[
(x - \zeta_{2^\beta})(x + \zeta_{2^\beta}) = x^2 - \zeta_{2^{\beta-1}}
\]
\[
(x - \zeta_{2^\beta})(x + i\zeta_{2^\beta}) = x^2 + (i - 1)\zeta_{2^\beta} - i\zeta_{2^{\beta-1}}
\]
\[
(x - \zeta_{2^\beta})(x - i\zeta_{2^\beta}) = x^2 - (i + 1)\zeta_{2^\beta} + i\zeta_{2^{\beta-1}}.
\]
But $\beta \geq 4$ so $i \in L(\zeta_{2^{\beta-2}})$ and the only way for any of these polynomials to be in $L(\zeta_{2^{\beta-2}})$ is for $\zeta_{2^{\beta-1}}$ to be in $L(\zeta_{2^{\beta-2}})$ as claimed. In fact if the total degree has to be 2 then the minimal polynomial is plainly the first of these.

**Proof of statement (B.1):** Suppose that $p^\alpha \mid J$ with $\alpha \geq 2$ if $p \neq 2$ and $\alpha \geq 3$ if $p = 2$. We first show that this implies:
\[
k(\zeta_J) = k(\zeta_{J/p}).
\]
If not we would have
\[
[k' \cap Q(\zeta) : k' \cap Q(\zeta_{J/p})] = \frac{\phi(J)/[k(\zeta) : k]}{\phi(J/p)/[k(\zeta_{J/p}) : k]} = \frac{p}{[k(\zeta) : k(\zeta_{J/p})]} = 1
\]
but this would imply that
\[
k' = k' \cap Q(\zeta) = k' \cap Q(\zeta_{J/p})
\]
contradicting the minimality of \(J\). Applying the lemma with \(L = k(\zeta_{J/p^n})\) we see at once that
\[
k(\zeta) = \begin{cases} k(\zeta_{J/p^n-1}) & p \neq 2 \\ k(\zeta_{J/2^n-2}) & p = 2. \end{cases}
\]
In fact we have
\[
k(\zeta) = k(\zeta_M).
\]
To see this we note that
\[
[k(\zeta_{M_{p^n-1}}) : k(\zeta_M)] = \frac{[k(\zeta) : k(\zeta_{J/p^n-1})][k(\zeta_{J/p^n-1}) : k(\zeta_M)]}{[k(\zeta) : k(\zeta_{M_{p^n-1}})]}
\]
\[
= \frac{[k(\zeta_{J/p^n-1}) : k(\zeta_M)]}{[k(\zeta) : k(\zeta_{M_{p^n-1}})]} \tag{B.2}
\]
where \(p^n \parallel J\) if \(p \neq 2\) and \(2^{n+1} \parallel J\) when \(p = 2\). Now the left-hand side of (B.2) is a power of \(p\):
\[
[k(\zeta_{M_{p^n-1}}) : k(\zeta_M)] = \frac{p^{n-1}}{[k \cap Q(\zeta_{M_{p^n-1}}) : k \cap Q(\zeta_M)]}
\]
while the right-hand side divides
\[
[k(\zeta_{J/p^n-1}) : k(\zeta_M)] = \frac{J/M p^{n-1}}{[k \cap Q(\zeta_{J/p^n-1}) : k \cap Q(\zeta_M)]}
\]
which is coprime to $p$. Hence in fact

$$k(\zeta_{M p^n-1}) = k(\zeta_M).$$

So $\zeta_{p^\alpha}$ is in $k(\zeta_M)$ for all $p^\alpha \parallel J$ and $k(\zeta_J) = k(\zeta_M)$ as claimed.

The relation (B.1) then follows easily from:

$$[k' : \mathbb{Q}] = [k' \cap \mathbb{Q}(\zeta_J) : \mathbb{Q}]$$

$$= [k' \cap \mathbb{Q}(\zeta_J) : k' \cap \mathbb{Q}(\zeta_M)][k' \cap \mathbb{Q}(\zeta_M) : \mathbb{Q}]$$

$$= \frac{\phi(J)/\phi(M)}{[k(\zeta_J) : k(\zeta_M)]}[k' \cap \mathbb{Q}(\zeta_M) : \mathbb{Q}]$$

$$= \frac{J}{M} [k' \cap \mathbb{Q}(\zeta_M) : \mathbb{Q}].$$
BIBLIOGRAPHY


VITA

Christopher Gerald Pinner was born in Great Yarmouth, England, on August 12, 1964, the son of Gerald and Margaret Pinner.

After leaving Gorleston Grammar School in 1982 he studied Mathematics at the University of York, receiving a First Class Bachelor of Science degree in June 1985. He then went on to New College, Oxford, to research number theory under the supervision of Dr. Roger Heath-Brown. He completed his Masters thesis "Topics in Analytic Number Theory" in August 1988 and was duly admitted to the degree of Master of Science (by research) by Oxford University in July 1989.

In September 1988 he came to the United States and entered the Graduate School of the University of Texas.

Permanent address: 4, Kennedy Ave,
Gorleston,
Gt. Yarmouth,
Norfolk NR31 6TA.
England.

This dissertation was typeset ¹ with \LaTeX \textsuperscript{X} by the author.

¹The \LaTeX \textsuperscript{X} document preparation system was developed by Leslie Lamport as a special version of Donald Knuth's \TeX program for computer typesetting. \TeX is a trademark of the American Mathematical Society. The \LaTeX \textsuperscript{X} macro package for The University of Texas at Austin dissertation format was written by Khe-Sing The.